(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 May 2001 (25.05.2001)

PCT

(10) International Publication Number
WO 01/36473 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/US00/31581

(22) International Filing Date: 16 November 2000 (16.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/165,838	16 November 1999 (16.11.1999)	US
60/166,071	17 November 1999 (17.11.1999)	US
60/166,678	19 November 1999 (19.11.1999)	US
60/173,396	28 December 1999 (28.12.1999)	US
60/184,129	22 February 2000 (22.02.2000)	US
60/185,421	28 February 2000 (28.02.2000)	US
60/185,554	28 February 2000 (28.02.2000)	US
60/186,530	2 March 2000 (02.03.2000)	US
60/186,811	3 March 2000 (03.03.2000)	US
60/188,114	9 March 2000 (09.03.2000)	US
60/190,310	17 March 2000 (17.03.2000)	US
60/190,800	21 March 2000 (21.03.2000)	US
60/198,568	20 April 2000 (20.04.2000)	US
60/201,190	2 May 2000 (02.05.2000)	US
60/203,111	8 May 2000 (08.05.2000)	US
60/207,094	25 May 2000 (25.05.2000)	US

(71) Applicant *(for all designated States except US)*: **PHARMACIA & UPJOHN COMPANY** [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **VOGELI, Gabriel** [US/US]; 3005 First Avenue, Seattle, WA 98121 (US). **WOOD, Linda, S.** [US/US]; 10193 Fox Hollow, Portage, MI 49024 (US). **PARODI, Luis, A.** [SE/SE]; Grevgatan 24, S-115 43 Stockholm (SE). **HIEBSCH, Ronald, R.** [US/US]; 917 Lane Boulevard, Kalamazoo, MI 49001 (US). **LIND, Peter** [SE/SE]; Borjegatan 31C, S-751 29 Uppsala (SE). **SLIGHTOM, Jerry** [US/US]; 3305 Lorraine Avenue, Kalamazoo, MI 49008 (US). **SCHELLIN, Kathleen, A.** [US/US]; 361 Jason Court, Portage, MI 49024 (US). **KAYTES, Paul, S.** [US/US]; 3506-D Shadow Bend Drive, Kalamazoo, MI 49004 (US). **BANNIGAN, Christopher, M.** [US/US]; 2411 Kopka Court, Bay City, MI 48708 (US). **RUFF, Valerie** [US/US]; 7301 Sandpiper Street, Portage, MI 49024 (US). **SEJLITZ, Torsten** [SE/SE]; Sankt Eriksgatan 59, S-112 34 Stockholm (SE). **HUFF, Rita, M.** [US/US]; 3627 B Avenue, West, Plainwell, MI 49080 (US).

(74) Agents: **DELUCA, Mark** et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, One Liberty Place, 46th Floor, Philadelphia, PA 19103 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GE, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
Without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/36473 A2

(54) **Title:** NOVEL G PROTEIN-COUPLED RECEPTORS

(57) **Abstract:** The present invention provides a gene encoding a G protein-coupled receptor termed nGPCR-x; constructs and recombinant host cells incorporating the genes; the nGPCR-x polypeptides encoded by the gene; antibodies to the nGPCR-x polypeptides; and methods of making and using all of the foregoing.

NOVEL G PROTEIN-COUPLED RECEPTORS

CROSS-REFERENCE TO RELATED APPLICATIONS

The present application claims priority of Application Serial No. 60/165,838, filed 1999 November 16; Serial No. 60/166,071, filed 1999 November 17; Serial No. 60/166,678 filed 1999 November 19; Serial No. 60/173,396, filed 1999 December 28; Serial No. 60/184,129, filed 2000 February 22; Serial No. 60/188,114, filed 2000 March 9; Serial No. 60/185,421, filed 2000 February 28; Serial No. 60/186,811, filed 2000 March 3; Serial No. 60/186,530, filed 2000 March 2; Serial No. 60/207,094, filed 2000 May 25; Serial No. 60/203,111, filed 2000 May 8; Serial No. 60/190,310, filed 2000 March 17; Serial No. 60/201,190, filed 2000 May 2; Serial No. 60/185,554, filed 2000 February 28; Serial No. 60/198,568, filed 2000 April 20; and Serial No. 60/190,800, filed 2000 March 21, each of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

The present invention relates generally to the fields of genetics and cellular and molecular biology. More particularly, the invention relates to novel G protein coupled receptors, to polynucleotides that encode such novel receptors, to reagents such as antibodies, probes, primers and kits comprising such antibodies, probes, primers related to the same, and to methods which use the novel G protein coupled receptors, polynucleotides or reagents.

BACKGROUND OF THE INVENTION

The G protein-coupled receptors (GPCRs) form a vast superfamily of cell surface receptors which are characterized by an amino-terminal extracellular domain, a carboxyl-terminal intracellular domain, and a serpentine structure that passes through the cell membrane seven times. Hence, such receptors are sometimes also referred to as seven transmembrane (7TM) receptors. These seven transmembrane domains define three extracellular loops and three intracellular loops, in addition to the amino- and carboxy- terminal domains. The extracellular portions of the receptor have a role in recognizing and binding one or more extracellular binding partners

(*e.g.*, ligands), whereas the intracellular portions have a role in recognizing and communicating with downstream molecules in the signal transduction cascade.

The G protein-coupled receptors bind a variety of ligands including calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and even photons, and are important in the normal (and sometimes the aberrant) function of many cell types. [See generally Strosberg, *Eur. J. Biochem.* 196:1-10 (1991) and Bohm *et al., Biochem J.* 322:1-18 (1997).] When a specific ligand binds to its corresponding receptor, the ligand typically stimulates the receptor to activate a specific heterotrimeric guanine-nucleotide-binding regulatory protein (G-protein) that is coupled to the intracellular portion of the receptor. The G protein in turn transmits a signal to an effector molecule within the cell, by either stimulating or inhibiting the activity of that effector molecule. These effector molecules include adenylate cyclase, phospholipases and ion channels. Adenylate cyclase and phospholipases are enzymes that are involved in the production of the second messenger molecules cAMP, inositol triphosphate and diacyglycerol. It is through this sequence of events that an extracellular ligand stimuli exerts intracellular changes through a G protein-coupled receptor. Each such receptor has its own characteristic primary structure, expression pattern, ligand-binding profile, and intracellular effector system.

Because of the vital role of G protein-coupled receptors in the communication between cells and their environment, such receptors are attractive targets for therapeutic intervention, for example by activating or antagonizing such receptors. For receptors having a known ligand, the identification of agonists or antagonists may be sought specifically to enhance or inhibit the action of the ligand. Some G protein-coupled receptors have roles in disease pathogenesis (*e.g.*, certain chemokine receptors that act as HIV co-receptors may have a role in AIDS pathogenesis), and are attractive targets for therapeutic intervention even in the absence of knowledge of the natural ligand of the receptor. Other receptors are attractive targets for therapeutic intervention by virtue of their expression pattern in tissues or cell types that are themselves attractive targets for therapeutic intervention. Examples of this latter category of receptors include receptors expressed in immune cells, which can be targeted to either inhibit autoimmune responses or to enhance immune responses to fight pathogens or cancer; and receptors expressed in the brain or other neural organs and tissues, which are likely targets in the treatment of schizophrenia, depression,

bipolar disease, or other neurological disorders. This latter category of receptor is also useful as a marker for identifying and/or purifying (*e.g.*, via fluorescence-activated cell sorting) cellular subtypes that express the receptor. Unfortunately, only a limited number of G protein receptors from the central nervous system (CNS) are known. Thus, a need exists for G protein-coupled receptors that have been identified and show promise as targets for therapeutic intervention in a variety of animals, including humans.

SUMMARY OF THE INVENTION

The present invention relates to an isolated nucleic acid molecule that comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a fragment thereof. The nucleic acid molecule encodes at least a portion of nGPCR-x. In some embodiments, the nucleic acid molecule comprises a sequence that encodes a polypeptide comprising even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a fragment thereof. In some embodiments, the nucleic acid molecule comprises a sequence homologous to odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a fragment thereof. In some embodiments, the nucleic acid molecule comprises a sequence selected from the group consisting of odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, and fragments thereof.

According to some embodiments, the present invention provides vectors which comprise the nucleic acid molecule of the invention. In some embodiments, the vector is an expression vector.

According to some embodiments, the present invention provides host cells which comprise the vectors of the invention. In some embodiments, the host cells comprise expression vectors.

The present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence complementary to at least a portion of a sequence from an odd numbered sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, said portion comprising at least 10 nucleotides.

The present invention provides a method of producing a polypeptide comprising a sequence from an even numbered sequence ranging from SEQ ID NO: 2

to SEQ ID NO: 94 and SEQ ID NO: 186, or a homolog or fragment thereof. The method comprising the steps of introducing a recombinant expression vector that includes a nucleotide sequence that encodes the polypeptide into a compatible host cell, growing the host cell under conditions for expression of the polypeptide and recovering the polypeptide.

The present invention provides an isolated antibody which binds to an epitope on a polypeptide comprising a sequence from an even numbered sequence ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a homolog or fragment thereof.

The present invention provides an method of inducing an immune response in a mammal against a polypeptide comprising a sequence from an even numbered sequence ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, or a homolog or fragment thereof. The method comprises administering to a mammal an amount of the polypeptide sufficient to induce said immune response.

The present invention provides a method for identifying a compound which binds nGPCR-x. The method comprises the steps of: contacting nGPCR-x with a compound and determining whether the compound binds nGPCR-x.

The present invention provides a method for identifying a compound which binds a nucleic acid molecule encoding nGPCR-x. The method comprises the steps of contacting said nucleic acid molecule encoding nGPCR-x with a compound and determining whether said compound binds said nucleic acid molecule.

The present invention provides a method for identifying a compound which modulates the activity of nGPCR-x. The method comprises the steps of contacting nGPCR-x with a compound and determining whether nGPCR-x activity has been modulated.

The present invention provides a method of identifying an animal homolog of nGPCR-x. The method comprises the steps screening a nucleic acid database of the animal with an odd numbered sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a portion thereof and determining whether a portion of said library or database is homologous to said odd numbered sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or portion thereof.

The present invention provides a method of identifying an animal homolog of nGPCR-x. The methods comprises the steps screening a nucleic acid library of the animal with a nucleic acid molecule having an odd numbered nucleotide sequence

ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a portion thereof; and determining whether a portion of said library or database is homologous to said odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or a portion thereof.

Another aspect of the present invention relates to methods of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor. The methods comprise the steps of assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering an amino acid sequence, expression, or biological activity of at least one nGPCR that is expressed in the brain. The nGPCR comprise an amino acid sequence selected from the group consisting of: SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof. A diagnosis of the disorder or predisposition is made from the presence or absence of the mutation. The presence of a mutation altering the amino acid sequence, expression, or biological activity of the nGPCR in the nucleic acid correlates with an increased risk of developing the disorder.

The present invention further relates to methods of screening for an nGPCR-40 or nGPCR-54 hereditary schizophrenia genotype in a human patient. The methods comprise the steps of providing a biological sample comprising nucleic acid from the patient, in which the nucleic acid includes sequences corresponding to allelles of nGPCR-40 or nGPCR-54. The presence of one or more mutations in the nGPCR-40 allelle or the nGPCR-54 allelle is detected indicative of a hereditary schizophrenia genotype.

The present invention provides kits for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor. The kits include an oligonucleotide useful as a probe for identifying polymorphisms in a human nGPCR-40 gene or a human nGPCR-54 gene. The oligonucleotide comprises 6-50 nucleotides in a sequence that is identical or complementary to a sequence of a wild type human nGPCR-40 or nGPCR-54 gene sequence or nGPCR-40 or nGPCR-54 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution. The kit also includes a media packaged with the oligonucleotide. The media contains information for identifying polymorphisms that correlate with schizophrenia or a

genetic predisposition therefor, the polymophisms being identifiable using the oligonucleotide as a probe.

The present invention further relates to methods of identifying nGPCR allelic variants that correlates with mental disorders. The methods comprise the steps of providing biological samples that comprise nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny, and detecting in the nucleic acid the presence of one or more mutations in an nGPCR that is expressed in the brain. The nGPCR comprises an amino acid sequence selected from the group consisting of SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof. The nucleic acid includes sequences corresponding to the gene or genes encoding nGPCR. The one or more mutations detected indicate an allelic variant that correlates with a mental disorder.

The present invention further relates to purified polynucleotides comprising nucleotide sequences encoding alleles of nGPCR-40 or nGPCR-54 from a human with schizophrenia. The polynucleotide hybridizes to the complement of SEQ ID NO:83 or of SEQ ID NO:85 under the following hybridization conditions: (a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate and (b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS. The polynucleotide that encodes nGPCR-40 or nGPCR-54 amino acid sequence of the human differs from SEQ ID NO:84 or SEQ ID NO:86 by at least one residue.

The present invention also provides methods for identifying a modulator of biological activity of nGPCR-40 or nGPCR-54 comprising the steps of contacting a cell that expresses nGPCR-40 or nGPCR-54 in the presence and in the absence of a putative modulator compound and measuring nGPCR-40 or nGPCR-54 biological activity in the cell. The decreased or increased nGPCR-40 or nGPCR-54 biological activity in the presence versus absence of the putative modulator is indicative of a modulator of biological activity.

The present invention further provides methods to identify compounds useful for the treatment of schizophrenia. The methods comprise the steps of contacting a composition comprising nGPCR-40 with a compound suspected of binding nGPCR-40 or contacting a composition comprising nGPCR-54 with a compound suspected of

binding nGPCR-54. The binding between nGPCR-40 and the compound suspected of binding nGPCR-40 or between nGPCR-54 and the compound suspected of binding nGPCR-54 is detected. Compounds identified as binding nGPCR-40 or nGPCR-54 are candidate compounds useful for the treatment of schizophrenia.

The present invention further provides methods for identifying a compound useful as a modulator of binding between nGPCR-40 and a binding partner of nGPCR-40 or between nGPCR-54 and a binding partner of nGPCR-54. The methods comprise the steps of contacting the binding partner and a composition comprising nGPCR-40 or nGPCR-54 in the presence and in the absence of a putative modulator compound and detecting binding between the binding partner and nGPCR-40 or nGPCR-54. Decreased or increased binding between the binding partner and nGPCR-40 or nGPCR-54 in the presence of the putative modulator, as compared to binding in the absence of the putative modulator is indicative a modulator compound useful for the treatment of schizophrenia.

Another aspect of the present invention relates to methods of purifying a G protein from a sample containing a G protein. The methods comprise the steps of contacting the sample with an nGPCR for a time sufficient to allow the G protein to form a complex with the nGPCR; isolating the complex from remaining components of the sample; maintaining the complex under conditions which result in dissociation of the G protein from the nGPCR; and isolating said G protein from the nGPCR.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Definitions

Various definitions are made throughout this document. Most words have the meaning that would be attributed to those words by one skilled in the art. Words specifically defined either below or elsewhere in this document have the meaning provided in the context of the present invention as a whole and as are typically understood by those skilled in the art.

"Synthesized" as used herein and understood in the art, refers to polynucleotides produced by purely chemical, as opposed to enzymatic, methods. "Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means.

By the term "region" is meant a physically contiguous portion of the primary structure of a biomolecule. In the case of proteins, a region is defined by a contiguous portion of the amino acid sequence of that protein.

The term "domain" is herein defined as referring to a structural part of a biomolecule that contributes to a known or suspected function of the biomolecule. Domains may be co-extensive with regions or portions thereof; domains may also incorporate a portion of a biomolecule that is distinct from a particular region, in addition to all or part of that region . Examples of GPCR protein domains include, but are not limited to, the extracellular (*i.e.*, N-terminal), transmembrane and cytoplasmic (*i.e.*, C-terminal) domains, which are co-extensive with like-named regions of GPCRs; each of the seven transmembrane segments of a GPCR; and each of the loop segments (both extracellular and intracellular loops) connecting adjacent transmembrane segments.

As used herein, the term "activity" refers to a variety of measurable indicia suggesting or revealing binding, either direct or indirect; affecting a response, i.e. having a measurable affect in response to some exposure or stimulus, including, for example, the affinity of a compound for directly binding a polypeptide or polynucleotide of the invention, or, for example, measurement of amounts of upstream or downstream proteins or other similar functions after some stimulus or event.

Unless indicated otherwise, as used herein, the abbreviation in lower case (gpcr) refers to a gene, cDNA, RNA or nucleic acid sequence, while the upper case version (GPCR) refers to a protein, polypeptide, peptide, oligopeptide, or amino acid sequence. The term "nGPCR-x" refers to any of the nGPCRs taught herein, while specific reference to a nGPCR (for example nGPCR-5) refers only to that specific nGPCR.

As used herein, the term "antibody" is meant to refer to complete, intact antibodies, and Fab, Fab', F(ab)2, and other fragments thereof. Complete, intact antibodies include monoclonal antibodies such as murine monoclonal antibodies, chimeric antibodies and humanized antibodies.

As used herein, the term "binding" means the physical or chemical interaction between two proteins or compounds or associated proteins or compounds or combinations thereof. Binding includes ionic, non-ionic, Hydrogen bonds, Van der Waals, hydrophobic interactions, etc. The physical interaction, the binding, can be

either direct or indirect, indirect being through or due to the effects of another protein or compound. Direct binding refers to interactions that do not take place through or due to the effect of another protein or compound but instead are without other substantial chemical intermediates. Binding may be detected in many different manners. As a non-limiting example, the physical binding interaction between a nGPCR-x of the invention and a compound can be detected using a labeled compound. Alternatively, functional evidence of binding can be detected using, for example, a cell transfected with and expressing a nGPCR-x of the invention. Binding of the transfected cell to a ligand of the nGPCR that was transfected into the cell provides functional evidence of binding. Other methods of detecting binding are well-known to those of skill in the art.

As used herein, the term "compound" means any identifiable chemical or molecule, including, but not limited to, small molecule, peptide, protein, sugar, nucleotide, or nucleic acid, and such compound can be natural or synthetic.

As used herein, the term "complementary" refers to Watson-Crick basepairing between nucleotide units of a nucleic acid molecule.

As used herein, the term "contacting" means bringing together, either directly or indirectly, a compound into physical proximity to a polypeptide or polynucleotide of the invention. The polypeptide or polynucleotide can be in any number of buffers, salts, solutions etc. Contacting includes, for example, placing the compound into a beaker, microtiter plate, cell culture flask, or a microarray, such as a gene chip, or the like, which contains the nucleic acid molecule, or polypeptide encoding the nGPCR or fragment thereof.

As used herein, the phrase "homologous nucleotide sequence," or "homologous amino acid sequence," or variations thereof, refers to sequences characterized by a homology, at the nucleotide level or amino acid level, of at least the specified percentage. Homologous nucleotide sequences include those sequences coding for isoforms of proteins. Such isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. Homologous nucleotide sequences include nucleotide sequences encoding for a protein of a species other than humans, including, but not limited to, mammals. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A homologous nucleotide sequence does

not, however, include the nucleotide sequence encoding other known GPCRs. Homologous amino acid sequences include those amino acid sequences which contain conservative amino acid substitutions and which polypeptides have the same binding and/or activity. A homologous amino acid sequence does not, however, include the amino acid sequence encoding other known GPCRs. Percent homology can be determined by, for example, the Gap program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, Madison WI), using the default settings, which uses the algorithm of Smith and Waterman (Adv. Appl. Math., 1981, 2, 482-489, which is incorporated herein by reference in its entirety).

As used herein, the term "isolated" nucleic acid molecule refers to a nucleic acid molecule (DNA or RNA) that has been removed from its native environment. Examples of isolated nucleic acid molecules include, but are not limited to, recombinant DNA molecules contained in a vector, recombinant DNA molecules maintained in a heterologous host cell, partially or substantially purified nucleic acid molecules, and synthetic DNA or RNA molecules.

As used herein, the terms "modulates" or "modifies" means an increase or decrease in the amount, quality, or effect of a particular activity or protein.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues which has a sufficient number of bases to be used in a polymerase chain reaction (PCR). This short sequence is based on (or designed from) a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue. Oligonucleotides comprise portions of a DNA sequence having at least about 10 nucleotides and as many as about 50 nucleotides, preferably about 15 to 30 nucleotides. They are chemically synthesized and may be used as probes.

As used herein, the term "probe" refers to nucleic acid sequences of variable length, preferably between at least about 10 and as many as about 6,000 nucleotides, depending on use. They are used in the detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are usually obtained from a natural or recombinant source, are highly specific and much slower to hybridize than oligomers. They may be single- or double-stranded and carefully designed to have specificity in PCR, hybridization membrane-based, or ELISA-like technologies.

The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, cell signaling, or cell survival. An abnormal condition may also include obesity, diabetic complications such as retinal degeneration, and irregularities in glucose uptake and metabolism, and fatty acid uptake and metabolism.

Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation.

Abnormal differentiation conditions include, but are not limited to, neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates. Abnormal cell signaling conditions include, but are not limited to, psychiatric disorders involving excess neurotransmitter activity.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or

treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

By "amplification" it is meant increased numbers of DNA or RNA in a cell compared with normal cells. "Amplification" as it refers to RNA can be the detectable presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1 to 2-fold, and preferably more, compared to the basal level.

As used herein, the phrase "stringent hybridization conditions" or "stringent conditions" refers to conditions under which a probe, primer, or oligonucleotide will hybridize to its target sequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present in excess, at T_m, 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (e.g. 10 to 50 nucleotides) and at least about 60°C for longer probes, primers or oligonucleotides.

Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

The amino acid sequences are presented in the amino to carboxy direction, from left to right. The amino and carboxy groups are not presented in the sequence. The nucleotide sequences are presented by single strand only, in the 5' to 3' direction, from left to right. Nucleotides and amino acids are represented in the manner recommended by the IUPAC-IUB Biochemical Nomenclature Commission or (for amino acids) by three letters code.

Polynucleotides

The present invention provides purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, both single- and double-stranded, including splice variants thereof) that encode unknown G protein-coupled receptors heretofore termed novel GPCRs, or nGPCRs. These genes are described herein and designated herein collectively as nGPCR-x (where x is 1, 3, 4, 5, 9, 11, 12, 14, 15, 18, 16, 17, 20, 21, 22, 24, 27, 28, 31, 32, 33, 34, 35, 36, 37, 38, 40, 41, 53, 54, 55, 56, 57, 58, 59, or 60). That is, these genes are described herein and designated herein as nGPCR-1 (also referred to as beGPCR-1), nGPCR-3 (also referred to as beGPCR-3), nGPCR-4 (also referred to as beGPCR-4), nGPCR-5 (also referred to as beGPCR-5 and TL-GPCR-5), nGPCR-9 (also referred to as beGPCR-9), nGPCR-11 (also referred to as beGPCR-11), nGPCR-12 (also referred to as beGPCR-12), nGPCR-14 (also referred to as beGPCR-14), nGPCR-15 (also referred to as beGPCR-15), nGPCR-18 (also referred to as beGPCR-18), nGPCR-16 (also referred to as beGPCR-16), nGPCR-17 (also referred to as beGPCR-17), nGPCR-20 (also referred to as beGPCR-20), nGPCR-21 (also referred to as beGPCR-21), nGPCR-22 (also referred to as beGPCR-22), nGPCR-24 (also referred to as beGPCR-24), nGPCR-27 (also referred to as beGPCR-27), nGPCR-28 (also referred to as beGPCR-28), nGPCR-31 (also referred to as beGPCR-31), nGPCR-32 (also referred to as beGPCR-32), nGPCR-33 (also referred to as beGPCR-33), nGPCR-34 (also referred to as beGPCR-34), nGPCR-35 (also referred to as beGPCR-35), nGPCR-36 (also referred to as beGPCR-36), nGPCR-37 (also referred to as beGPCR-37), nGPCR-38 (also referred to as beGPCR-38), nGPCR-40 (also referred to as beGPCR-40), nGPCR-41 (also referred to as beGPCR-41), nGPCR-53, nGPCR-54, nGPCR-55, nGPCR-56, nGPCR-57, nGPCR-58, nGPCR-59, and

nGPCR-60. Table 1 below identifies the novel gene sequence nGPCR-x designation, the SEQ ID NO: of the gene sequence, the SEQ ID NO: of the polypeptide encoded thereby, and the U.S. Provisional Application in which the gene sequence has been disclosed.

Table 1

nGPCR	Nucleotide Sequence (SEQ ID NO:)	Amino acid Sequence (SEQ ID NO:)	Originally filed in:		nGPCR	Nucleotide Sequence (SEQ ID NO:)	Amino acid Sequence (SEQ ID NO:)	Originally filed in:
1	1	2	A		32	39	40	B
1	73	74	E		33	41	42	C
3	3	4	A		34	43	44	C
3	185	186	P		35	45	46	C
4	5	6	A		36	47	48	C
5	7	8	A		37	49	50	C
5	75	76	F		38	51	52	C
9	9	10	A		40	53	54	C
9	77	78	G		40	83	84	J
11	11	12	A		41	55	56	C
11	79	80	H		53	57	58	D
12	13	14	A		54	59	60	D
14	15	16	A		54	85	86	K
15	17	18	A		55	61	62	D
18	19	20	A		56	63	64	D
16	21	22	B		56	87	88	L
16	81	82	I		56	89	90	M
17	23	24	B		57	65	66	D
20	25	26	B		58	67	68	D
21	27	28	B		58	91	92	N
22	29	30	B		58	93	94	O
24	31	32	B		59	69	70	D
27	33	34	B		60	71	72	D
28	35	36	B					
31	37	38	B					

Legend

A= Ser. No. 60/165,838
B= Ser. No. 60/166,071
C= Ser. No. 60/166,678
D= Ser. No. 60/173,396
E= Ser. No. 60/184,129
F= Ser. No. 60/188,114
G= Ser. No. 60/185,421
H= Ser. No. 60/186,811
I= Ser. No. 60/186,530
J= Ser. No. 60/207,094
K= Ser. No. 60/203,111
L= Ser. No. 60/190,310
M= Ser. No. 60/201,190
N= Ser. No. 60/185554
O= Ser. No. 60/190,800
P= Ser. No. 60/198,568

When a specific nGPCR is identified (for example nGPCR-5), it is understood that only that specific nGPCR is being referred to.

As described in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEQ ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence

SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO: 86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, amino acid sequence SEQ ID NO:68, SEQ ID NO: 92, SEQ ID NO:94) and nGPCR-3 (nucleic acid sequence SEQ ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors.

The invention provides purified and isolated polynucleotides (*e.g.*, cDNA, genomic DNA, synthetic DNA, RNA, or combinations thereof, whether single- or double-stranded) that comprise a nucleotide sequence encoding the amino acid sequence of the polypeptides of the invention. Such polynucleotides are useful for recombinantly expressing the receptor and also for detecting expression of the receptor in cells (*e.g.*, using Northern hybridization and *in situ* hybridization assays). Such polynucleotides also are useful in the design of antisense and other molecules for the suppression of the expression of nGPCR-x in a cultured cell, a tissue, or an animal; for therapeutic purposes; or to provide a model for diseases or conditions characterized by aberrant nGPCR-x expression. Specifically excluded from the definition of polynucleotides of the invention are entire isolated, non-recombinant native chromosomes of host cells. A preferred polynucleotide has the sequence of the sequence set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, which correspond to naturally occurring nGPCR-x sequences. It will be appreciated that numerous other polynucleotide sequences exist that also encode nGPCR-x having the sequence set forth in even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, due to the well-known degeneracy of the universal genetic code.

The invention also provides a purified and isolated polynucleotide comprising a nucleotide sequence that encodes a mammalian polypeptide, wherein the polynucleotide hybridizes to a polynucleotide having the sequence set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID

NO: 185 or the non-coding strand complementary thereto, under the following hybridization conditions:

(a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate; and

(b) washing 2 times for 30 minutes each at 60°C in a wash solution comprising 0.1% SSC, 1% SDS. Polynucleotides that encode a human allelic variant are highly preferred.

The present invention relates to molecules which comprise the gene sequences that encode the nGPCRs; constructs and recombinant host cells incorporating the gene sequences; the novel GPCR polypeptides encoded by the gene sequences; antibodies to the polypeptides and homologs; kits employing the polynucleotides and polypeptides, and methods of making and using all of the foregoing. In addition, the present invention relates to homologs of the gene sequences and of the polypeptides and methods of making and using the same.

Genomic DNA of the invention comprises the protein-coding region for a polypeptide of the invention and is also intended to include allelic variants thereof. It is widely understood that, for many genes, genomic DNA is transcribed into RNA transcripts that undergo one or more splicing events wherein intron (*i.e.*, non-coding regions) of the transcripts are removed, or "spliced out." RNA transcripts that can be spliced by alternative mechanisms, and therefore be subject to removal of different RNA sequences but still encode a nGPCR-x polypeptide, are referred to in the art as splice variants which are embraced by the invention. Splice variants comprehended by the invention therefore are encoded by the same original genomic DNA sequences but arise from distinct mRNA transcripts. Allelic variants are modified forms of a wild-type gene sequence, the modification resulting from recombination during chromosomal segregation or exposure to conditions which give rise to genetic mutation. Allelic variants, like wild type genes, are naturally occurring sequences (as opposed to non-naturally occurring variants that arise from *in vitro* manipulation).

The invention also comprehends cDNA that is obtained through reverse transcription of an RNA polynucleotide encoding nGPCR-x (conventionally followed by second strand synthesis of a complementary strand to provide a double-stranded DNA).

Preferred DNA sequences encoding human nGPCR-x polypeptides are set out in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. A preferred DNA of the invention comprises a double stranded molecule along with the complementary molecule (the "non-coding strand" or "complement") having a sequence unambiguously deducible from the coding strand according to Watson-Crick base-pairing rules for DNA. Also preferred are other polynucleotides encoding the nGPCR-x polypeptide of even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, which differ in sequence from the polynucleotides of odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, by virtue of the well-known degeneracy of the universal nuclear genetic code.

The invention further embraces other species, preferably mammalian, homologs of the human nGPCR-x DNA. Species homologs, sometimes referred to as "orthologs," in general, share at least 35%, at least 40%, at least 45%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% homology with human DNA of the invention. Generally, percent sequence "homology" with respect to polynucleotides of the invention may be calculated as the percentage of nucleotide bases in the candidate sequence that are identical to nucleotides in the nGPCR-x sequence set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity.

Polynucleotides of the invention permit identification and isolation of polynucleotides encoding related nGPCR-x polypeptides, such as human allelic variants and species homologs, by well-known techniques including Southern and/or Northern hybridization, and polymerase chain reaction (PCR). Examples of related polynucleotides include human and non-human genomic sequences, including allelic variants, as well as polynucleotides encoding polypeptides homologous to nGPCR-x and structurally related polypeptides sharing one or more biological, immunological, and/or physical properties of nGPCR-x. Non-human species genes encoding proteins homologous to nGPCR-x can also be identified by Southern and/or PCR analysis and are useful in animal models for nGPCR-x disorders. Knowledge of the sequence of a human nGPCR-x DNA also makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences

encoding nGPCR-x expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express nGPCR-x. Polynucleotides of the invention may also provide a basis for diagnostic methods useful for identifying a genetic alteration(s) in a nGPCR-x locus that underlies a disease state or states, which information is useful both for diagnosis and for selection of therapeutic strategies.

According to the present invention, the nGPCR-x nucleotide sequences disclosed herein may be used to identify homologs of the nGPCR-x, in other animals, including but not limited to humans and other mammals, and invertebrates. Any of the nucleotide sequences disclosed herein, or any portion thereof, can be used, for example, as probes to screen databases or nucleic acid libraries, such as, for example, genomic or cDNA libraries, to identify homologs, using screening procedures well known to those skilled in the art. Accordingly, homologs having at least 50%, more preferably at least 60%, more preferably at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, and most preferably at least 100% homology with nGPCR-x sequences can be identified.

The disclosure herein of full-length polynucleotides encoding nGPCR-x polypeptides makes readily available to the worker of ordinary skill in the art every possible fragment of the full-length polynucleotide.

One preferred embodiment of the present invention provides an isolated nucleic acid molecule comprising a sequence homologous to odd numbered sequences selected from the group consisting of SEQ ID NO:1 to SEQ ID NO:93, SEQ ID NO: 185, and fragments thereof. Another preferred embodiment provides an isolated nucleic acid molecule comprising a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO: 93, SEQ ID NO: 185 and fragments thereof.

As used in the present invention, fragments of nGPCR-x-encoding polynucleotides comprise at least 10, and preferably at least 12, 14, 16, 18, 20, 25, 50, or 75 consecutive nucleotides of a polynucleotide encoding nGPCR-x. Preferably, fragment polynucleotides of the invention comprise sequences unique to the nGPCR-x-encoding polynucleotide sequence, and therefore hybridize under highly stringent or moderately stringent conditions only (*i.e.*, "specifically") to polynucleotides encoding nGPCR-x (or fragments thereof). Polynucleotide fragments of genomic sequences of

the invention comprise not only sequences unique to the coding region, but also include fragments of the full-length sequence derived from introns, regulatory regions, and/or other non-translated sequences. Sequences unique to polynucleotides of the invention are recognizable through sequence comparison to other known polynucleotides, and can be identified through use of alignment programs routinely utilized in the art, *e.g.*, those made available in public sequence databases. Such sequences also are recognizable from Southern hybridization analyses to determine the number of fragments of genomic DNA to which a polynucleotide will hybridize. Polynucleotides of the invention can be labeled in a manner that permits their detection, including radioactive, fluorescent, and enzymatic labeling.

Fragment polynucleotides are particularly useful as probes for detection of full-length or fragments of nGPCR-x polynucleotides. One or more polynucleotides can be included in kits that are used to detect the presence of a polynucleotide encoding nGPCR-x, or used to detect variations in a polynucleotide sequence encoding nGPCR-x.

The invention also embraces DNAs encoding nGPCR-x polypeptides that hybridize under moderately stringent or high stringency conditions to the non-coding strand, or complement, of the polynucleotides set forth in odd numbered sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185.

Exemplary highly stringent hybridization conditions are as follows: hybridization at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% Dextran sulfate, and washing twice for 30 minutes at 60°C in a wash solution comprising 0.1 X SSC and 1% SDS. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described Ausubel *et al.* (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10. Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook, *et al.*, (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

With the knowledge of the nucleotide sequence information disclosed in the present invention, one skilled in the art can identify and obtain nucleotide sequences

which encode nGPCR-x from different sources (*i.e.*, different tissues or different organisms) through a variety of means well known to the skilled artisan and as disclosed by, for example, Sambrook et al., "Molecular cloning: a laboratory manual", Second Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), which is incorporated herein by reference in its entirety.

For example, DNA that encodes nGPCR-x may be obtained by screening of mRNA, cDNA, or genomic DNA with oligonucleotide probes generated from the nGPCR-x gene sequence information provided herein. Probes may be labeled with a detectable group, such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with procedures known to the skilled artisan and used in conventional hybridization assays, as described by, for example, Sambrook *et al.*

A nucleic acid molecule comprising any of the nGPCR-x nucleotide sequences described above can alternatively be synthesized by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers produced from the nucleotide sequences provided herein. See U.S. Patent Numbers 4,683,195 to Mullis *et al.* and 4,683,202 to Mullis. The PCR reaction provides a method for selectively increasing the concentration of a particular nucleic acid sequence even when that sequence has not been previously purified and is present only in a single copy in a particular sample. The method can be used to amplify either single- or double-stranded DNA. The essence of the method involves the use of two oligonucleotide probes to serve as primers for the template-dependent, polymerase mediated replication of a desired nucleic acid molecule.

A wide variety of alternative cloning and *in vitro* amplification methodologies are well known to those skilled in the art. Examples of these techniques are found in, for example, Berger *et al.*, *Guide to Molecular Cloning Techniques*, Methods in Enzymology 152, Academic Press, Inc., San Diego, CA (Berger), which is incorporated herein by reference in its entirety.

Automated sequencing methods can be used to obtain or verify the nucleotide sequence of nGPCR-x. The nGPCR-x nucleotide sequences of the present invention are believed to be 100% accurate. However, as is known in the art, nucleotide sequence obtained by automated methods may contain some errors. Nucleotide sequences determined by automation are typically at least about 90%, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence

of a given nucleic acid molecule. The actual sequence may be more precisely determined using manual sequencing methods, which are well known in the art. An error in a sequence which results in an insertion or deletion of one or more nucleotides may result in a frame shift in translation such that the predicted amino acid sequence will differ from that which would be predicted from the actual nucleotide sequence of the nucleic acid molecule, starting at the point of the mutation.

The nucleic acid molecules of the present invention, and fragments derived therefrom, are useful for screening for restriction fragment length polymorphism (RFLP) associated with certain disorders, as well as for genetic mapping.

The polynucleotide sequence information provided by the invention makes possible large-scale expression of the encoded polypeptide by techniques well known and routinely practiced in the art.

Vectors

Another aspect of the present invention is directed to vectors, or recombinant expression vectors, comprising any of the nucleic acid molecules described above. Vectors are used herein either to amplify DNA or RNA encoding nGPCR-x and/or to express DNA which encodes nGPCR-x. Preferred vectors include, but are not limited to, plasmids, phages, cosmids, episomes, viral particles or viruses, and integratable DNA fragments (*i.e.*, fragments integratable into the host genome by homologous recombination). Preferred viral particles include, but are not limited to, adenoviruses, baculoviruses, parvoviruses, herpesviruses, poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses. Preferred expression vectors include, but are not limited to, pcDNA3 (Invitrogen) and pSVL (Pharmacia Biotech). Other expression vectors include, but are not limited to, pSPORT™ vectors, pGEM™ vectors (Promega), pPROEXvectors™ (LTI, Bethesda, MD), Bluescript™ vectors (Stratagene), pQE™ vectors (Qiagen), pSE420™ (Invitrogen), and pYES2™(Invitrogen).

Expression constructs preferably comprise GPCR-x-encoding polynucleotides operatively linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator. Expression control DNA sequences include promoters, enhancers, operators, and regulatory element binding sites generally, and are typically selected based on the expression systems in which the expression construct is to be utilized. Preferred promoter and enhancer sequences are generally selected for the ability to increase gene expression, while operator sequences are generally selected

for the ability to regulate gene expression. Expression constructs of the invention may also include sequences encoding one or more selectable markers that permit identification of host cells bearing the construct. Expression constructs may also include sequences that facilitate, and preferably promote, homologous recombination in a host cell. Preferred constructs of the invention also include sequences necessary for replication in a host cell.

Expression constructs are preferably utilized for production of an encoded protein, but may also be utilized simply to amplify a nGPCR-x-encoding polynucleotide sequence. In preferred embodiments, the vector is an expression vector wherein the polynucleotide of the invention is operatively linked to a polynucleotide comprising an expression control sequence. Autonomously replicating recombinant expression constructs such as plasmid and viral DNA vectors incorporating polynucleotides of the invention are also provided. Preferred expression vectors are replicable DNA constructs in which a DNA sequence encoding nGPCR-x is operably linked or connected to suitable control sequences capable of effecting the expression of the nGPCR-x in a suitable host. DNA regions are operably linked or connected when they are functionally related to each other. For example, a promoter is operably linked or connected to a coding sequence if it controls the transcription of the sequence. Amplification vectors do not require expression control domains, but rather need only the ability to replicate in a host, usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants. The need for control sequences in the expression vector will vary depending upon the host selected and the transformation method chosen. Generally, control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding and sequences which control the termination of transcription and translation.

Preferred vectors preferably contain a promoter that is recognized by the host organism. The promoter sequences of the present invention may be prokaryotic, eukaryotic or viral. Examples of suitable prokaryotic sequences include the P_R and P_L promoters of bacteriophage lambda (The bacteriophage Lambda, Hershey, A. D., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1973), which is incorporated herein by reference in its entirety; Lambda II, Hendrix, R. W., Ed., Cold Spring Harbor Press, Cold Spring Harbor, NY (1980), which is incorporated herein by reference in its entirety); the trp, recA, heat shock, and lacZ promoters of *E. coli* and

the SV40 early promoter (Benoist *et al. Nature*, 1981, *290*, 304-310, which is incorporated herein by reference in its entirety). Additional promoters include, but are not limited to, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, Rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine, and human metalothionein.

Additional regulatory sequences can also be included in preferred vectors. Preferred examples of suitable regulatory sequences are represented by the Shine-Dalgarno of the replicase gene of the phage MS-2 and of the gene cII of bacteriophage lambda. The Shine-Dalgarno sequence may be directly followed by DNA encoding nGPCR-x and result in the expression of the mature nGPCR-x protein.

Moreover, suitable expression vectors can include an appropriate marker that allows the screening of the transformed host cells. The transformation of the selected host is carried out using any one of the various techniques well known to the expert in the art and described in Sambrook *et al., supra*.

An origin of replication can also be provided either by construction of the vector to include an exogenous origin or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter may be sufficient. Alternatively, rather than using vectors which contain viral origins of replication, one skilled in the art can transform mammalian cells by the method of co-transformation with a selectable marker and nGPCR-x DNA. An example of a suitable marker is dihydrofolate reductase (DHFR) or thymidine kinase (*see*, U.S. Patent No. 4,399,216).

Nucleotide sequences encoding GPCR-x may be recombined with vector DNA in accordance with conventional techniques, including blunt-ended or staggered-ended termini for ligation, restriction enzyme digestion to provide appropriate termini, filling in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesiderable joining, and ligation with appropriate ligases. Techniques for such manipulation are disclosed by Sambrook et al., *supra* and are well known in the art. Methods for construction of mammalian expression vectors are disclosed in, for example, Okayama *et al., Mol. Cell. Biol.*, **1983**, *3*, 280, Cosman *et al., Mol. Immunol.*, **1986**, *23*, 935, Cosman *et al., Nature*, **1984**, *312*, 768, EP-A-0367566, and WO 91/18982, each of which is incorporated herein by reference in its entirety.

Host cells

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, comprising a polynucleotide of the invention (or vector of the invention) in a manner that permits expression of the encoded nGPCR-x polypeptide. Polynucleotides of the invention may be introduced into the host cell as part of a circular plasmid, or as linear DNA comprising an isolated protein coding region or a viral vector. Methods for introducing DNA into the host cell that are well known and routinely practiced in the art include transformation, transfection, electroporation, nuclear injection, or fusion with carriers such as liposomes, micelles, ghost cells, and protoplasts. Expression systems of the invention include bacterial, yeast, fungal, plant, insect, invertebrate, vertebrate, and mammalian cells systems.

The invention provides host cells that are transformed or transfected (stably or transiently) with polynucleotides of the invention or vectors of the invention. As stated above, such host cells are useful for amplifying the polynucleotides and also for expressing the nGPCR-x polypeptide or fragment thereof encoded by the polynucleotide.

In still another related embodiment, the invention provides a method for producing a nGPCR-x polypeptide (or fragment thereof) comprising the steps of growing a host cell of the invention in a nutrient medium and isolating the polypeptide or variant thereof from the cell or the medium. Because nGPCR-x is a seven transmembrane receptor, it will be appreciated that, for some applications, such as certain activity assays, the preferable isolation may involve isolation of cell membranes containing the polypeptide embedded therein, whereas for other applications a more complete isolation may be preferable.

According to some aspects of the present invention, transformed host cells having an expression vector comprising any of the nucleic acid molecules described above are provided. Expression of the nucleotide sequence occurs when the expression vector is introduced into an appropriate host cell. Suitable host cells for expression of the polypeptides of the invention include, but are not limited to, prokaryotes, yeast, and eukaryotes. If a prokaryotic expression vector is employed, then the appropriate host cell would be any prokaryotic cell capable of expressing the cloned sequences. Suitable prokaryotic cells include, but are not limited to, bacteria

of the genera *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*, *Streptomyces*, and *Staphylococcus*.

If an eukaryotic expression vector is employed, then the appropriate host cell would be any eukaryotic cell capable of expressing the cloned sequence. Preferably, eukaryotic cells are cells of higher eukaryotes. Suitable eukaryotic cells include, but are not limited to, non-human mammalian tissue culture cells and human tissue culture cells. Preferred host cells include, but are not limited to, insect cells, HeLa cells, Chinese hamster ovary cells (CHO cells), African green monkey kidney cells (COS cells), human 293 cells, and murine 3T3 fibroblasts. Propagation of such cells in cell culture has become a routine procedure (*see*, Tissue Culture, Academic Press, Kruse and Patterson, eds. (1973), which is incorporated herein by reference in its entirety).

In addition, a yeast host may be employed as a host cell. Preferred yeast cells include, but are not limited to, the genera *Saccharomyces*, *Pichia*, and *Kluveromyces*. Preferred yeast hosts are *S. cerevisiae* and *P. pastoris*. Preferred yeast vectors can contain an origin of replication sequence from a 2T yeast plasmid, an autonomously replication sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Shuttle vectors for replication in both yeast and *E. coli* are also included herein.

Alternatively, insect cells may be used as host cells. In a preferred embodiment, the polypeptides of the invention are expressed using a baculovirus expression system (*see*, Luckow *et al.*, *Bio/Technology*, **1988**, *6*, 47, Baculovirus Expression Vectors: A Laboratory Manual, O'Rielly *et al.* (Eds.), W.H. Freeman and Company, New York, **1992**, and U.S. Patent No. 4,879,236, each of which is incorporated herein by reference in its entirety). In addition, the MAXBAC™ complete baculovirus expression system (Invitrogen) can, for example, be used for production in insect cells.

Host cells of the invention are a valuable source of immunogen for development of antibodies specifically immunoreactive with nGPCR-x. Host cells of the invention are also useful in methods for the large-scale production of nGPCR-x polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells, or from the medium in which the cells are grown, by purification methods known in the art, *e.g.*, conventional chromatographic methods including immunoaffinity chromatography, receptor

affinity chromatography, hydrophobic interaction chromatography, lectin affinity chromatography, size exclusion filtration, cation or anion exchange chromatography, high pressure liquid chromatography (HPLC), reverse phase HPLC, and the like. Still other methods of purification include those methods wherein the desired protein is expressed and purified as a fusion protein having a specific tag, label, or chelating moiety that is recognized by a specific binding partner or agent. The purified protein can be cleaved to yield the desired protein, or can be left as an intact fusion protein. Cleavage of the fusion component may produce a form of the desired protein having additional amino acid residues as a result of the cleavage process.

Knowledge of nGPCR-x DNA sequences allows for modification of cells to permit, or increase, expression of endogenous nGPCR-x. Cells can be modified (*e.g.*, by homologous recombination) to provide increased expression by replacing, in whole or in part, the naturally occurring nGPCR-x promoter with all or part of a heterologous promoter so that the cells express nGPCR-x at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to endogenous nGPCR-x encoding sequences. (See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No.WO 92/20808, and PCT International Publication No. WO 91/09955.) It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamoyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the nGPCR-x coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the nGPCR-x coding sequences in the cells.

Knock-outs

The DNA sequence information provided by the present invention also makes possible the development (*e.g.*, by homologous recombination or "knock-out" strategies; see Capecchi, *Science 244*:1288-1292 (1989), which is incorporated herein by reference) of animals that fail to express functional nGPCR-x or that express a variant of nGPCR-x. Such animals (especially small laboratory animals such as rats, rabbits, and mice) are useful as models for studying the *in vivo* activities of nGPCR-x and modulators of nGPCR-x.

Antisense

Also made available by the invention are anti-sense polynucleotides that recognize and hybridize to polynucleotides encoding nGPCR-x. Full-length and fragment anti-sense polynucleotides are provided. Fragment antisense molecules of the invention include (i) those that specifically recognize and hybridize to nGPCR-x RNA (as determined by sequence comparison of DNA encoding nGPCR-x to DNA encoding other known molecules). Identification of sequences unique to nGPCR-x encoding polynucleotides can be deduced through use of any publicly available sequence database, and/or through use of commercially available sequence comparison programs. After identification of the desired sequences, isolation through restriction digestion or amplification using any of the various polymerase chain reaction techniques well known in the art can be performed. Anti-sense polynucleotides are particularly relevant to regulating expression of nGPCR-x by those cells expressing nGPCR-x mRNA.

Antisense nucleic acids (preferably 10 to 30 base-pair oligonucleotides) capable of specifically binding to nGPCR-x expression control sequences or nGPCR-x RNA are introduced into cells (*e.g.*, by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the nGPCR-x target nucleotide sequence in the cell and prevents transcription and/or translation of the target sequence. Phosphorothioate and methylphosphonate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by adding poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5' end. Suppression of nGPCR-x expression at either the transcriptional or translational level is useful to generate cellular or animal models for diseases/conditions characterized by aberrant nGPCR-x expression.

Antisense oligonucleotides, or fragments of odd numbered nucleotide sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 or sequences complementary or homologous thereto, derived from the nucleotide sequences of the present invention encoding nGPCR-x are useful as diagnostic tools for probing gene expression in various tissues. For example, tissue can be probed *in situ* with oligonucleotide probes carrying detectable groups by conventional autoradiography techniques to investigate native expression of this enzyme or pathological conditions relating thereto. Antisense oligonucleotides are preferably

directed to regulatory regions of odd numbered nucleotide sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185, or mRNA corresponding thereto, including, but not limited to, the initiation codon, TATA box, enhancer sequences, and the like.

Transcription factors

The nGPCR-x sequences taught in the present invention facilitate the design of novel transcription factors for modulating nGPCR-x expression in native cells and animals, and cells transformed or transfected with nGPCR-x polynucleotides. For example, the Cys_2-His_2 zinc finger proteins, which bind DNA via their zinc finger domains, have been shown to be amenable to structural changes that lead to the recognition of different target sequences. These artificial zinc finger proteins recognize specific target sites with high affinity and low dissociation constants, and are able to act as gene switches to modulate gene expression. Knowledge of the particular nGPCR-x target sequence of the present invention facilitates the engineering of zinc finger proteins specific for the target sequence using known methods such as a combination of structure-based modeling and screening of phage display libraries (Segal *et al.*, Proc. Natl. Acad. Sci. (USA) 96:2758-2763 (1999); Liu *et al.*, Proc. Natl. Acad. Sci. (USA) 94:5525-5530 (1997); Greisman *et al.*, Science 275:657-661 (1997); Choo *et al.*, J. Mol. Biol. 273:525-532 (1997)). Each zinc finger domain usually recognizes three or more base pairs. Since a recognition sequence of 18 base pairs is generally sufficient in length to render it unique in any known genome, a zinc finger protein consisting of 6 tandem repeats of zinc fingers would be expected to ensure specificity for a particular sequence (Segal *et al.*) The artificial zinc finger repeats, designed based on nGPCR-x sequences, are fused to activation or repression domains to promote or suppress nGPCR-x expression (Liu *et al.)* Alternatively, the zinc finger domains can be fused to the TATA box-binding factor (TBP) with varying lengths of linker region between the zinc finger peptide and the TBP to create either transcriptional activators or repressors (Kim *et al.*, Proc. Natl. Acad. Sci. (USA) 94:3616-3620 (1997). Such proteins and polynucleotides that encode them, have utility for modulating nGPCR-x expression *in vivo* in both native cells, animals and humans; and/or cells transfected with nGPCR-x-encoding sequences. The novel transcription factor can be delivered to the target cells by transfecting constructs that express the transcription factor (gene therapy), or by introducing the protein. Engineered zinc finger proteins can also be designed to bind

RNA sequences for use in therapeutics as alternatives to antisense or catalytic RNA methods (McColl *et al.*, Proc. Natl. Acad. Sci. (USA) 96:9521-9526 (1997); Wu *et al.*, Proc. Natl. Acad. Sci. (USA) 92:344-348 (1995)). The present invention contemplates methods of designing such transcription factors based on the gene sequence of the invention, as well as customized zinc finger proteins, that are useful to modulate nGPCR-x expression in cells (native or transformed) whose genetic complement includes these sequences.

Polypeptides

The invention also provides purified and isolated mammalian nGPCR-x polypeptides encoded by a polynucleotide of the invention. Presently preferred is a human nGPCR-x polypeptide comprising the amino acid sequence set out in even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186 or fragments thereof comprising an epitope specific to the polypeptide. By "epitope specific to" is meant a portion of the nGPCR receptor that is recognizable by an antibody that is specific for the nGPCR, as defined in detail below.

Although the sequences provided are particular human sequences, the invention is intended to include within its scope other human allelic variants; non-human mammalian forms of nGPCR-x, and other vertebrate forms of nGPCR-x.

It will be appreciated that extracellular epitopes are particularly useful for generating and screening for antibodies and other binding compounds that bind to receptors such as nGPCR-x. Thus, in another preferred embodiment, the invention provides a purified and isolated polypeptide comprising at least one extracellular domain (*e.g.*, the N-terminal extracellular domain or one of the three extracellular loops) of nGPCR-x. Purified and isolated polypeptides comprising the N-terminal extracellular domain of nGPCR-x are highly preferred. Also preferred is a purified and isolated polypeptide comprising a nGPCR-x fragment selected from the group consisting of the N-terminal extracellular domain of nGPCR-x, transmembrane domains of nGPCR-x, an extracellular loop connecting transmembrane domains of nGPCR-x, an intracellular loop connecting transmembrane domains of nGPCR-x, the C-terminal cytoplasmic region of nGPCR-x, and fusions thereof. Such fragments may be continuous portions of the native receptor. However, it will also be appreciated that knowledge of the nGPCR-x gene and protein sequences as provided herein permits recombining of various domains that are not contiguous in the native protein. Using a FORTRAN computer program called "tmtrest.all" [Parodi *et al.*,

Comput. Appl. Biosci. 5:527-535 (1994)], nGPCR-x was shown to contain transmembrane-spanning domains.

The invention also embraces polypeptides that have at least 99%, at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55% or at least 50% identity and/or homology to the preferred polypeptide of the invention. Percent amino acid sequence "identity" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the nGPCR-x sequence after aligning both sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Percent sequence "homology" with respect to the preferred polypeptide of the invention is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the residues in the nGPCR-x sequence after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and also considering any conservative substitutions as part of the sequence identity.

In one aspect, percent homology is calculated as the percentage of amino acid residues in the smaller of two sequences which align with identical amino acid residue in the sequence being compared, when four gaps in a length of 100 amino acids may be introduced to maximize alignment [Dayhoff, in Atlas of Protein Sequence and Structure, Vol. 5, p. 124, National Biochemical Research Foundation, Washington, D.C. (1972), incorporated herein by reference].

Polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. Use of mammalian host cells is expected to provide for such post-translational modifications (*e.g.*, glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Glycosylated and non-glycosylated forms of nGPCR-x polypeptides are embraced by the invention.

The invention also embraces variant (or analog) nGPCR-x polypeptides. In one example, insertion variants are provided wherein one or more amino acid residues supplement a nGPCR-x amino acid sequence. Insertions may be located at either or both termini of the protein, or may be positioned within internal regions of the nGPCR-x amino acid sequence. Insertional variants with additional residues at either

or both termini can include, for example, fusion proteins and proteins including amino acid tags or labels.

Insertion variants include nGPCR-x polypeptides wherein one or more amino acid residues are added to a nGPCR-x acid sequence or to a biologically active fragment thereof.

Variant products of the invention also include mature nGPCR-x products, *i.e.*, nGPCR-x products wherein leader or signal sequences are removed, with additional amino terminal residues. The additional amino terminal residues may be derived from another protein, or may include one or more residues that are not identifiable as being derived from specific proteins. nGPCR-x products with an additional methionine residue at position -1 (Met^{-1}-nGPCR-x) are contemplated, as are variants with additional methionine and lysine residues at positions -2 and -1 (Met^{-2}-Lys^{-1}-nGPCR-x). Variants of nGPCR-x with additional Met, Met-Lys, Lys residues (or one or more basic residues in general) are particularly useful for enhanced recombinant protein production in bacterial host cells.

The invention also embraces nGPCR-x variants having additional amino acid residues that result from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide as part of a glutathione-S-transferase (GST) fusion product provides the desired polypeptide having an additional glycine residue at position -1 after cleavage of the GST component from the desired polypeptide. Variants that result from expression in other vector systems are also contemplated.

Insertional variants also include fusion proteins wherein the amino terminus and/or the carboxy terminus of nGPCR-x is/are fused to another polypeptide.

In another aspect, the invention provides deletion variants wherein one or more amino acid residues in a nGPCR-x polypeptide are removed. Deletions can be effected at one or both termini of the nGPCR-x polypeptide, or with removal of one or more non-terminal amino acid residues of nGPCR-x. Deletion variants, therefore, include all fragments of a nGPCR-x polypeptide.

The invention also embraces polypeptide fragments of the even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, wherein the fragments maintain biological (*e.g.*, ligand binding and/or intracellular signaling) immunological properties of a nGPCR-x polypeptide.

In one preferred embodiment of the invention, an isolated nucleic acid molecule comprises a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to even numbered sequences selected from the group consisting of: SEQ ID NO:2 to SEQ ID NO:94, SEQ ID NO: 186, and fragments thereof, wherein the nucleic acid molecule encoding at least a portion of nGPCR-x. In a more preferred embodiment, the isolated nucleic acid molecule comprises a sequence that encodes a polypeptide comprising even numbered sequences selected from the group consisting of SEQ ID NO:2 to SEQ ID NO: 94, SEQ ID NO: 186, and fragments thereof.

As used in the present invention, polypeptide fragments comprise at least 5, 10, 15, 20, 25, 30, 35, or 40 consecutive amino acids of the even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186. Preferred polypeptide fragments display antigenic properties unique to, or specific for, human nGPCR-x and its allelic and species homologs. Fragments of the invention having the desired biological and immunological properties can be prepared by any of the methods well known and routinely practiced in the art.

In still another aspect, the invention provides substitution variants of nGPCR-x polypeptides. Substitution variants include those polypeptides wherein one or more amino acid residues of a nGPCR-x polypeptide are removed and replaced with alternative residues. In one aspect, the substitutions are conservative in nature; however, the invention embraces substitutions that are also non-conservative. Conservative substitutions for this purpose may be defined as set out in Tables 2, 3, or 4 below.

Variant polypeptides include those wherein conservative substitutions have been introduced by modification of polynucleotides encoding polypeptides of the invention. Amino acids can be classified according to physical properties and contribution to secondary and tertiary protein structure. A conservative substitution is recognized in the art as a substitution of one amino acid for another amino acid that has similar properties. Exemplary conservative substitutions are set out in Table 2 (from WO 97/09433, page 10, published March 13, 1997 (PCT/GB96/02197, filed 9/6/96), immediately below.

Table 2

Conservative Substitutions I

SIDE CHAIN CHARACTERISTIC	**AMINO ACID**
Aliphatic	
Non-polar	G A P
	I L V
Polar - uncharged	C S T M
	N Q
Polar - charged	D E
	K R
Aromatic	H F W Y
Other	N Q D E

Alternatively, conservative amino acids can be grouped as described in Lehninger, [Biochemistry, Second Edition; Worth Publishers, Inc. NY, NY (1975), pp.71-77] as set out in Table 3, below.

Table 3
Conservative Substitutions II

SIDE CHAIN CHARACTERISTIC	**AMINO ACID**
Non-polar (hydrophobic)	
A. Aliphatic:	A L I V P
B. Aromatic:	F W
C. Sulfur-containing:	M
D. Borderline:	G
Uncharged-polar	
A. Hydroxyl:	S T Y
B. Amides:	N Q
C. Sulfhydryl:	C
D. Borderline:	G
Positively Charged (Basic):	K R H
Negatively Charged (Acidic):	D E

As still another alternative, exemplary conservative substitutions are set out in Table 4, below.

Table 4
Conservative Substitutions III

Original Residue	**Exemplary Substitution**
Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe,

Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met (M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

It should be understood that the definition of polypeptides of the invention is intended to include polypeptides bearing modifications other than insertion, deletion, or substitution of amino acid residues. By way of example, the modifications may be covalent in nature, and include for example, chemical bonding with polymers, lipids, other organic, and inorganic moieties. Such derivatives may be prepared to increase circulating half-life of a polypeptide, or may be designed to improve the targeting capacity of the polypeptide for desired cells, tissues, or organs. Similarly, the invention further embraces nGPCR-x polypeptides that have been covalently modified to include one or more water-soluble polymer attachments such as polyethylene glycol, polyoxyethylene glycol, or polypropylene glycol. Variants that display ligand binding properties of native nGPCR-x and are expressed at higher levels, as well as variants that provide for constitutively active receptors, are particularly useful in assays of the invention; the variants are also useful in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant nGPCR-x activity.

In a related embodiment, the present invention provides compositions comprising purified polypeptides of the invention. Preferred compositions comprise, in addition to the polypeptide of the invention, a pharmaceutically acceptable (*i.e.*, sterile and non-toxic) liquid, semisolid, or solid diluent that serves as a pharmaceutical vehicle, excipient, or medium. Any diluent known in the art may be used. Exemplary diluents include, but are not limited to, water, saline solutions, polyoxyethylene sorbitan monolaurate, magnesium stearate, methyl- and propylhydroxybenzoate, talc, alginates, starches, lactose, sucrose, dextrose, sorbitol, mannitol, glycerol, calcium phosphate, mineral oil, and cocoa butter.

Variants that display ligand binding properties of native nGPCR-x and are expressed at higher levels, as well as variants that provide for constitutively active

receptors, are particularly useful in assays of the invention; the variants are also useful in assays of the invention and in providing cellular, tissue and animal models of diseases/conditions characterized by aberrant nGPCR-x activity.

The G protein-coupled receptor functions through a specific heterotrimeric guanine-nucleotide-binding regulatory protein (G-protein) coupled to the intracellular portion of the G protein-coupled receptor molecule. Accordingly, the G protein-coupled receptor has a specific affinity to G protein. G proteins specifically bind to guanine nucleotides. Isolation of G proteins provides a means to isolate guanine nucleotides. G Proteins may be isolated using commercially available anti-G protein antibodies or isolated G protein-coupled receptors. Similarly, G proteins may be detected in a sample isolated using commercially available detectable anti-G protein antibodies or isolated G protein-coupled receptors.

According to the present invention, the isolated n-GPCR-x proteins of the present invention are useful to isolate and purify G proteins from samples such as cell lysates. Example 15 below sets forth an example of isolation of G proteins using isolated n-GPCR-x proteins. Such methodolgy may be used in place of the use of commercially available anti-G protein antibodies which are used to isolate G proteins. Moreover, G proteins may be detected using n-GPCR-x proteins in place of commercially available detectable anti-G protein antibodies. Since n-GPCR-x proteins specifically bind to G proteins, they can be employed in any specific use where G protein specific affinity is required such as those uses where commercially available anti-G protein antibodies are employed.

Antibodies

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR sequences which specifically recognize a polypeptide of the invention) specific for nGPCR-x or fragments thereof. Preferred antibodies of the invention are human antibodies that are produced and identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab′, $F(ab')_2$, and F_v, are also provided by the invention. The term "specific for," when used to describe antibodies of the invention, indicates that the variable regions of the antibodies of the

invention recognize and bind nGPCR-x polypeptides exclusively (*i.e.*, are able to distinguish nGPCR-x polypeptides from other known GPCR polypeptides by virtue of measurable differences in binding affinity, despite the possible existence of localized sequence identity, homology, or similarity between nGPCR-x and such polypeptides). It will be understood that specific antibodies may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and, in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* (Eds.), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the nGPCR-x polypeptides of the invention are also contemplated, provided that the antibodies are specific for nGPCR-x polypeptides. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

The invention provides an antibody that is specific for the nGPCR-x of the invention. Antibody specificity is described in greater detail below. However, it should be emphasized that antibodies that can be generated from polypeptides that have previously been described in the literature and that are capable of fortuitously cross-reacting with nGPCR-x (*e.g.*, due to the fortuitous existence of a similar epitope in both polypeptides) are considered "cross-reactive" antibodies. Such cross-reactive antibodies are not antibodies that are "specific" for nGPCR-x. The determination of whether an antibody is specific for nGPCR-x or is cross-reactive with another known receptor is made using any of several assays, such as Western blotting assays, that are well known in the art. For identifying cells that express nGPCR-x and also for modulating nGPCR-x-ligand binding activity, antibodies that specifically bind to an extracellular epitope of the nGPCR-x are preferred.

In one preferred variation, the invention provides monoclonal antibodies. Hybridomas that produce such antibodies also are intended as aspects of the invention. In yet another variation, the invention provides a humanized antibody. Humanized antibodies are useful for *in vivo* therapeutic indications.

In another variation, the invention provides a cell-free composition comprising polyclonal antibodies, wherein at least one of the antibodies is an antibody of the invention specific for nGPCR-x. Antisera isolated from an animal is an exemplary

composition, as is a composition comprising an antibody fraction of an antisera that has been resuspended in water or in another diluent, excipient, or carrier.

In still another related embodiment, the invention provides an anti-idiotypic antibody specific for an antibody that is specific for nGPCR-x.

It is well known that antibodies contain relatively small antigen binding domains that can be isolated chemically or by recombinant techniques. Such domains are useful nGPCR-x binding molecules themselves, and also may be reintroduced into human antibodies, or fused to toxins or other polypeptides. Thus, in still another embodiment, the invention provides a polypeptide comprising a fragment of a nGPCR-x-specific antibody, wherein the fragment and the polypeptide bind to the nGPCR-x. By way of non-limiting example, the invention provides polypeptides that are single chain antibodies and CDR-grafted antibodies.

Non-human antibodies may be humanized by any of the methods known in the art. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

Antibodies of the invention are useful for, *e.g.*, therapeutic purposes (by modulating activity of nGPCR-x), diagnostic purposes to detect or quantitate nGPCR-x, and purification of nGPCR-x. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a control antigen for which the antibody is immunospecific.

Compositions

Mutations in the nGPCR-x gene that result in loss of normal function of the nGPCR-x gene product underlie nGPCR-x-related human disease states. The invention comprehends gene therapy to restore nGPCR-x activity to treat those disease states. Delivery of a functional nGPCR-x gene to appropriate cells is effected *ex vivo*, *in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). See, for example, Anderson, *Nature*, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, *Science*, *244*: 1275-1281 (1989); Verma, *Scientific American*: 68-84 (1990); and Miller, *Nature*, *357*: 455-460 (1992). Alternatively, it is contemplated that in other human disease states, preventing the expression of, or inhibiting the activity of, nGPCR-x will be useful in treating disease

states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of nGPCR-x.

Another aspect of the present invention is directed to compositions, including pharmaceutical compositions, comprising any of the nucleic acid molecules or recombinant expression vectors described above and an acceptable carrier or diluent. Preferably, the carrier or diluent is pharmaceutically acceptable. Suitable carriers are described in the most recent edition of *Remington's Pharmaceutical Sciences*, A. Osol, a standard reference text in this field, which is incorporated herein by reference in its entirety. Preferred examples of such carriers or diluents include, but are not limited to, water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Liposomes and nonaqueous vehicles such as fixed oils may also be used. The formulations are sterilized by commonly used techniques.

Also within the scope of the invention are compositions comprising polypeptides, polynucleotides, or antibodies of the invention that have been formulated with, *e.g.*, a pharmaceutically acceptable carrier.

The invention also provides methods of using antibodies of the invention. For example, the invention provides a method for modulating ligand binding of a nGPCR-x comprising the step of contacting the nGPCR-x with an antibody specific for the nGPCR-x, under conditions wherein the antibody binds the receptor.

GPCRs that may be expressed in the brain, such as nGPCR-x, provide an indication that aberrant nGPCR-x signaling activity may correlate with one or more neurological or psychological disorders. The invention also provides a method for treating a neurological or psychiatric disorder comprising the step of administering to a mammal in need of such treatment an amount of an antibody-like polypeptide of the invention that is sufficient to modulate ligand binding to a nGPCR-x in neurons of the mammal. nGPCR-x may also be expressed in other tissues, including but not limited to, peripheral blood lymphocytes, pancreas, ovary, uterus, testis, salivary gland, thyroid gland, kidney, adrenal gland, liver, bone marrow, prostate, fetal liver, colon, muscle, and fetal brain, and may be found in many other tissues. Within the brain, nGPCR-x mRNA transcripts may be found in many tissues, including, but not limited to, frontal lobe, hypothalamus, pons, cerebellum, caudate nucleus, and medulla. Tissues and brain regions where specific nGPCRs of the present invention are expressed are identified in the Examples below.

Kits

The present invention is also directed to kits, including pharmaceutical kits. The kits can comprise any of the nucleic acid molecules described above, any of the polypeptides described above, or any antibody which binds to a polypeptide of the invention as described above, as well as a negative control. The kit preferably comprises additional components, such as, for example, instructions, solid support, reagents helpful for quantification, and the like.

In another aspect, the invention features methods for detection of a polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a polypeptide having the sequence of even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, said probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease.

In preferred embodiments of the invention, the disease is selected from the group consisting of thyroid disorders (*e.g.* thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (*e.g.*, Crohn's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (*e.g.*, pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, *etc.*); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (*e.g.*, type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, *etc.*); proliferative diseases and cancers (*e.g.*, different cancers such as breast, colon, lung, *etc.*, and hyperproliferative disorders such as psoriasis, prostate hyperplasia, *etc.*); hormonal disorders (*e.g.*, male/female hormonal replacement, polycystic ovarian syndrome, alopecia, *etc.*); and sexual dysfunction, among others.

As described above and in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEQ ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO: 86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:67, SEQ ID NO 91, SEQ ID NO:93, amino acid sequence SEQ ID NO:68, SEQ ID NO: 92, SEQ ID NO:94) and nGPCR-3 (nucleic acid sequence SEQ ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors. Kits may be designed to detect either expression of polynucleotides encoding these proteins or the proteins themselves in order to identify tissue as being neurological. For example, oligonucleotide hybridization kits can be provided which include a container having an oligonucleotide probe specific for the n-GPCR-x-specific DNA and optionally, containers with positive and negative controls and/or instructions. Similarly, PCR kits can be provided which include a container having primers specific for the n-GPCR-x-specific sequences, DNA and optionally, containers with size markers, positive and negative controls and/or instructions.

Hybridization conditions should be such that hybridization occurs only with the genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined supra.

The diseases for which detection of genes in a sample could be diagnostic include diseases in which nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells. By "amplification" is meant increased numbers of DNA or RNA in a cell compared with normal cells.

The diseases that could be diagnosed by detection of nucleic acid in a sample preferably include central nervous system and metabolic diseases. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

Alternatively, immunoassay kits can be provided which have containers container having antibodies specific for the n-GPCR-x-protein and optionally, containers with positive and negative controls and/or instructions.

Kits may also be provided useful in the identification of GPCR binding partners such as natural ligands or modulators (agonists or antagonists). Substances useful for treatment of disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question. Substances that modulate the activity of the polypeptides preferably include, but are not limited to, antisense oligonucleotides, agonists and antagonists, and inhibitors of protein kinases.

Methods of inducing immune response

Another aspect of the present invention is directed to methods of inducing an immune response in a mammal against a polypeptide of the invention by administering to the mammal an amount of the polypeptide sufficient to induce an immune response. The amount will be dependent on the animal species, size of the animal, and the like but can be determined by those skilled in the art.

Methods of identifying ligands

The invention also provides assays to identify compounds that bind nGPCR-x. One such assay comprises the steps of: (a) contacting a composition comprising a nGPCR-x with a compound suspected of binding nGPCR-x; and (b) measuring binding between the compound and nGPCR-x. In one variation, the composition comprises a cell expressing nGPCR-x on its surface. In another variation, isolated nGPCR-x or cell membranes comprising nGPCR-x are employed. The binding may be measured directly, *e.g.*, by using a labeled compound, or may be measured indirectly by several techniques, including measuring intracellular signaling of

nGPCR-x induced by the compound (or measuring changes in the level of nGPCR-x signaling).

Specific binding molecules, including natural ligands and synthetic compounds, can be identified or developed using isolated or recombinant nGPCR-x products, nGPCR-x variants, or preferably, cells expressing such products. Binding partners are useful for purifying nGPCR-x products and detection or quantification of nGPCR-x products in fluid and tissue samples using known immunological procedures. Binding molecules are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of nGPCR-x, especially those activities involved in signal transduction.

The DNA and amino acid sequence information provided by the present invention also makes possible identification of binding partner compounds with which a nGPCR-x polypeptide or polynucleotide will interact. Methods to identify binding partner compounds include solution assays, *in vitro* assays wherein nGPCR-x polypeptides are immobilized, and cell-based assays. Identification of binding partner compounds of nGPCR-x polypeptides provides candidates for therapeutic or prophylactic intervention in pathologies associated with nGPCR-x normal and aberrant biological activity.

The invention includes several assay systems for identifying nGPCR-x binding partners. In solution assays, methods of the invention comprise the steps of (a) contacting a nGPCR-x polypeptide with one or more candidate binding partner compounds and (b) identifying the compounds that bind to the nGPCR-x polypeptide. Identification of the compounds that bind the nGPCR-x polypeptide can be achieved by isolating the nGPCR-x polypeptide/binding partner complex, and separating the binding partner compound from the nGPCR-x polypeptide. An additional step of characterizing the physical, biological, and/or biochemical properties of the binding partner compound is also comprehended in another embodiment of the invention. In one aspect, the nGPCR-x polypeptide/binding partner complex is isolated using an antibody immunospecific for either the nGPCR-x polypeptide or the candidate binding partner compound.

In still other embodiments, either the nGPCR-x polypeptide or the candidate binding partner compound comprises a label or tag that facilitates its isolation, and methods of the invention to identify binding partner compounds include a step of isolating the nGPCR-x polypeptide/binding partner complex through interaction with

the label or tag. An exemplary tag of this type is a poly-histidine sequence, generally around six histidine residues, that permits isolation of a compound so labeled using nickel chelation. Other labels and tags, such as the FLAG® tag (Eastman Kodak, Rochester, NY), well known and routinely used in the art, are embraced by the invention.

In one variation of an *in vitro* assay, the invention provides a method comprising the steps of (a) contacting an immobilized nGPCR-x polypeptide with a candidate binding partner compound and (b) detecting binding of the candidate compound to the nGPCR-x polypeptide. In an alternative embodiment, the candidate binding partner compound is immobilized and binding of nGPCR-x is detected. Immobilization is accomplished using any of the methods well known in the art, including covalent bonding to a support, a bead, or a chromatographic resin, as well as non-covalent, high affinity interactions such as antibody binding, or use of streptavidin/biotin binding wherein the immobilized compound includes a biotin moiety. Detection of binding can be accomplished (i) using a radioactive label on the compound that is not immobilized, (ii) using of a fluorescent label on the non-immobilized compound, (iii) using an antibody immunospecific for the non-immobilized compound, (iv) using a label on the non-immobilized compound that excites a fluorescent support to which the immobilized compound is attached, as well as other techniques well known and routinely practiced in the art.

The invention also provides cell-based assays to identify binding partner compounds of a nGPCR-x polypeptide. In one embodiment, the invention provides a method comprising the steps of contacting a nGPCR-x polypeptide expressed on the surface of a cell with a candidate binding partner compound and detecting binding of the candidate binding partner compound to the nGPCR-x polypeptide. In a preferred embodiment, the detection comprises detecting a calcium flux or other physiological event in the cell caused by the binding of the molecule.

Another aspect of the present invention is directed to methods of identifying compounds that bind to either nGPCR-x or nucleic acid molecules encoding nGPCR-x, comprising contacting nGPCR-x, or a nucleic acid molecule encoding the same, with a compound, and determining whether the compound binds nGPCR-x or a nucleic acid molecule encoding the same. Binding can be determined by binding assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression

cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, *Current Protocols in Molecular Biology*, **1999**, John Wiley & Sons, NY, which is incorporated herein by reference in its entirety. The compounds to be screened include (which may include compounds which are suspected to bind nGPCR-x, or a nucleic acid molecule encoding the same), but are not limited to, extracellular, intracellular, biologic or chemical origin. The methods of the invention also embrace ligands, especially neuropeptides, that are attached to a label, such as a radiolabel (*e.g.*, ^{125}I, ^{35}S, ^{32}P, ^{33}P, ^{3}H), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. Modulators falling within the scope of the invention include, but are not limited to, non-peptide molecules such as non-peptide mimetics, non-peptide allosteric effectors, and peptides. The nGPCR-x polypeptide or polynucleotide employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface or located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between nGPCR-x and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between nGPCR-x and its substrate caused by the compound being tested.

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to nGPCR-x is employed. Briefly, large numbers of different small peptide test compounds are synthesized on a solid substrate. The peptide test compounds are contacted with nGPCR-x and washed. Bound nGPCR-x is then detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Generally, an expressed nGPCR-x can be used for HTS binding assays in conjunction with its defined ligand, in this case the corresponding neuropeptide that activates it. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I, ^{3}H, ^{35}S or ^{32}P, by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable fluorescent derivative (Baindur *et al.*, *Drug Dev. Res.*, **1994**, *33*, 373-398; Rogers, *Drug Discovery Today*, **1997**, *2*, 156-160). Radioactive ligand specifically

bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate bound ligand from unbound ligand (Williams, *Med. Res. Rev.*, **1991**, *11*, 147-184; Sweetnam *et al.*, *J. Natural Products*, **1993**, *56*, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama, *Cur. Opinion Drug Disc. Dev.*, **1998**, *1*, 85-91 Bossé *et al.*, *J. Biomolecular Screening*, **1998**, *3*, 285-292.). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers, *Drug Discovery Today*, **1997**, *2*, 156-160; Hill, *Cur. Opinion* Drug Disc. *Dev.*, **1998**, *1*, 92-97).

Other assays may be used to identify specific ligands of a nGPCR-x receptor, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, Nature, 340:245-246 (1989), and Fields *et al.*, Trends in Genetics, 10:286-292 (1994), both of which are incorporated herein by reference. The two-hybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain,

cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a GPCR gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal.

The function of nGPCR-x gene products is unclear and no ligands have yet been found which bind the gene product. The yeast two-hybrid assay can also be used to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a nGPCR-x receptor, or fragment thereof, a fusion polynucleotide encoding both a nGPCR-x receptor (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein-coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (i.e., when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method that distinguishes between the folded and unfolded states of the target protein. The function of the

target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.*, Anal. Chem., 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with nGPCR-x. Radiolabeled competitive binding studies are described in A.H. Lin *et al. Antimicrobial Agents and Chemotherapy*, **1997**, vol. 41, no. 10. pp. 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

As described above and in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEQ ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO: 86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ

ID NO:67, SEQ ID NO:91, SEQ ID NO:93, amino acid sequence SEQ ID NO:68, SEQ ID NO: 92, SEQ ID NO:94), and nGPCR-3 (nucleic acid sequence SEQ ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors. Accordingly, natural binding partners of these molecules include neurotransmitters.

Identification of modulating agents

The invention also provides methods for identifying a modulator of binding between a nGPCR-x and a nGPCR-x binding partner, comprising the steps of: (a) contacting a nGPCR-x binding partner and a composition comprising a nGPCR-x in the presence and in the absence of a putative modulator compound; (b) detecting binding between the binding partner and the nGPCR-x; and (c) identifying a putative modulator compound or a modulator compound in view of decreased or increased binding between the binding partner and the nGPCR-x in the presence of the putative modulator, as compared to binding in the absence of the putative modulator.

nGPCR-x binding partners that stimulate nGPCR-x activity are useful as agonists in disease states or conditions characterized by insufficient nGPCR-x signaling (*e.g.*, as a result of insufficient activity of a nGPCR-x ligand). nGPCR-x binding partners that block ligand-mediated nGPCR-x signaling are useful as nGPCR-x antagonists to treat disease states or conditions characterized by excessive nGPCR-x signaling. In addition nGPCR-x modulators in general, as well as nGPCR-x polynucleotides and polypeptides, are useful in diagnostic assays for such diseases or conditions.

In another aspect, the invention provides methods for treating a disease or abnormal condition by administering to a patient in need of such treatment a substance that modulates the activity or expression of a polypeptide having the sequence of even numbered sequences ranging from SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186.

Agents that modulate (*i.e.*, increase, decrease, or block) nGPCR-x activity or expression may be identified by incubating a putative modulator with a cell containing a nGPCR-x polypeptide or polynucleotide and determining the effect of the putative modulator on nGPCR-x activity or expression. The selectivity of a compound that modulates the activity of nGPCR-x can be evaluated by comparing its effects on nGPCR-x to its effect on other GPCR compounds. Selective modulators

may include, for example, antibodies and other proteins, peptides, or organic molecules that specifically bind to a nGPCR-x polypeptide or a nGPCR-x-encoding nucleic acid. Modulators of nGPCR-x activity will be therapeutically useful in treatment of diseases and physiological conditions in which normal or aberrant nGPCR-x activity is involved. nGPCR-x polynucleotides, polypeptides, and modulators may be used in the treatment of such diseases and conditions as infections, such as viral infections caused by HIV-1 or HIV-2; pain; cancers; Parkinson's disease; hypotension; hypertension; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome, among others. nGPCR-x polynucleotides and polypeptides, as well as nGPCR-x modulators, may also be used in diagnostic assays for such diseases or conditions.

Methods of the invention to identify modulators include variations on any of the methods described above to identify binding partner compounds, the variations including techniques wherein a binding partner compound has been identified and the binding assay is carried out in the presence and absence of a candidate modulator. A modulator is identified in those instances where binding between the nGPCR-x polypeptide and the binding partner compound changes in the presence of the candidate modulator compared to binding in the absence of the candidate modulator compound. A modulator that increases binding between the nGPCR-x polypeptide and the binding partner compound is described as an enhancer or activator, and a modulator that decreases binding between the nGPCR-x polypeptide and the binding partner compound is described as an inhibitor.

The invention also comprehends high-throughput screening (HTS) assays to identify compounds that interact with or inhibit biological activity (*i.e.*, affect enzymatic activity, binding activity, *etc.*) of a nGPCR-x polypeptide. HTS assays permit screening of large numbers of compounds in an efficient manner. Cell-based HTS systems are contemplated to investigate nGPCR-x receptor-ligand interaction. HTS assays are designed to identify "hits" or "lead compounds" having the desired property, from which modifications can be designed to improve the desired property. Chemical modification of the "hit" or "lead compound" is often based on an identifiable structure/activity relationship between the "hit" and the nGPCR-x polypeptide.

Another aspect of the present invention is directed to methods of identifying compounds which modulate (*i.e.*, increase or decrease) activity of nGPCR-x comprising contacting nGPCR-x with a compound, and determining whether the compound modifies activity of nGPCR-x. The activity in the presence of the test compared is measured to the activity in the absence of the test compound. Where the activity of the sample containing the test compound is higher than the activity in the sample lacking the test compound, the compound will have increased activity. Similarly, where the activity of the sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound will have inhibited activity.

The present invention is particularly useful for screening compounds by using nGPCR-x in any of a variety of drug screening techniques. The compounds to be screened include (which may include compounds which are suspected to modulate nGPCR-x activity), but are not limited to, extracellular, intracellular, biologic or chemical origin. The nGPCR-x polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between nGPCR-x and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between nGPCR-x and its substrate caused by the compound being tested.

The activity of nGPCR-x polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesized peptide ligands. Alternatively, the activity of nGPCR-x polypeptides can be assayed by examining their ability to bind calcium ions, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and photons. Alternatively, the activity of the nGPCR-x polypeptides can be determined by examining the activity of effector molecules including, but not limited to, adenylate cyclase, phospholipases and ion channels. Thus, modulators of nGPCR-x polypeptide activity may alter a GPCR receptor function, such as a binding property of a receptor or an activity such as G protein-mediated signal transduction or membrane localization. In various embodiments of the method, the assay may take the form of an ion flux assay, a yeast growth assay, a non-hydrolyzable GTP assay such as a [^{35}S]-GTP S assay, a cAMP assay, an inositol triphosphate assay, a diacylglycerol assay, an Aequorin assay, a Luciferase assay, a FLIPR assay for intracellular Ca^{2+}

concentration, a mitogenesis assay, a MAP Kinase activity assay, an arachidonic acid release assay (*e.g.*, using [^{3}H]-arachidonic acid), and an assay for extracellular acidification rates, as well as other binding or function-based assays of nGPCR-x activity that are generally known in the art. In several of these embodiments, the invention comprehends the inclusion of any of the G proteins known in the art, such as G_{16}, G_{15}, or chimeric G_{qd5}, G_{qs5}, G_{qo5}, G_{q25}, and the like. nGPCR-x activity can be determined by methodologies that are used to assay for FaRP activity, which is well known to those skilled in the art. Biological activities of nGPCR-x receptors according to the invention include, but are not limited to, the binding of a natural or an unnatural ligand, as well as any one of the functional activities of GPCRs known in the art. Non-limiting examples of GPCR activities include transmembrane signaling of various forms, which may involve G protein association and/or the exertion of an influence over G protein binding of various guanidylate nucleotides; another exemplary activity of GPCRs is the binding of accessory proteins or polypeptides that differ from known G proteins.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into non-peptide mimetics of natural GPCR receptor ligands, peptide and non-peptide allosteric effectors of GPCR receptors, and peptides that may function as activators or inhibitors (competitive, uncompetitive and non-competitive) (*e.g.*, antibody products) of GPCR receptors. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries. Examples of peptide modulators of GPCR receptors exhibit the following primary structures: GLGPRPLRFamide, GNSFLRFamide, GGPQGPLRFamide, GPSGPLRFamide, PDVDHVFLRFamide, and pyro-EDVDHVFLRFamide.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, *Enzyme Assays: A Practical Approach*, eds. R. Eisenthal and M. J. Danson, 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

The use of cDNAs encoding GPCRs in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-

throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use of radiolabelled ligands in HTS binding assays for drug discovery (see Williams, *Medicinal Research Reviews*, **1991**, *11*, 147-184.; Sweetnam, *et al.*, *J. Natural Products*, **1993**, *56*, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see Hodgson, *Bio/Technology*, **1992**, *10*, 973-980; each of which is incorporated herein by reference in its entirety).

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg, *et al.*, *Trends in Pharmacological Sciences*, **1992**, *13*, 95-98), yeast (Pausch, *Trends in Biotechnology*, **1997**, *15*, 487-494), several kinds of insect cells (Vanden Broeck, *Int. Rev. Cytology*, **1996**, *164*, 189-268), amphibian cells (Jayawickreme *et al.*, *Current Opinion in Biotechnology*, **1997**, *8*, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt, *et al.*, *Eur. J. Pharmacology*, **1997**, *334*, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

In preferred embodiments of the invention, methods of screening for compounds that modulate nGPCR-x activity comprise contacting test compounds with nGPCR-x and assaying for the presence of a complex between the compound and nGPCR-x. In such assays, the ligand is typically labeled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular compound to bind to nGPCR-x.

It is well known that activation of heterologous receptors expressed in recombinant systems results in a variety of biological responses, which are mediated by G proteins expressed in the host cells. Occupation of a GPCR by an agonist results in exchange of bound GDP for GTP at a binding site on the G_α subunit; one can use a radioactive, non-hydrolyzable derivative of GTP, GTPγ[^{35}S], to measure binding of an agonist to the receptor (Sim *et al.*, Neuroreport, **1996**, *7*, 729-733). One can also use this binding to measure the ability of antagonists to bind to the receptor by decreasing binding of GTPγ[^{35}S] in the presence of a known agonist. One could

therefore construct a HTS based on GTPγ[^{35}S] binding, though this is not the preferred method.

The G proteins required for functional expression of heterologous GPCRs can be native constituents of the host cell or can be introduced through well-known recombinant technology. The G proteins can be intact or chimeric. Often, a nearly universally competent G protein (e.g., $G_{\alpha 16}$) is used to couple any given receptor to a detectable response pathway. G protein activation results in the stimulation or inhibition of other native proteins, events that can be linked to a measurable response.

Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch, *Trends in Biotechnology*, **1997**, *15*, 487-494); changes in intracellular Ca^{2+} concentration as measured by fluorescent dyes (Murphy, *et al.*, *Cur. Opinion Drug Disc. Dev.*, **1998**, *1*, 192-199). Fluorescence changes can also be used to monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder, *et al.*, *J. Biomolecular Screening*, **1996**, *1*, 75-80). Melanophores prepared from *Xenopus laevis* show a ligand-dependent change in pigment organization in response to heterologous GPCR activation; this response is adaptable to HTS formats (Jayawickreme *et al.*, *Cur. Opinion Biotechnology*, **1997**, *8*, 629-634). Assays are also available for the measurement of common second messengers, including cAMP, phosphoinositides and arachidonic acid, but these are not generally preferred for HTS.

Preferred methods of HTS employing these receptors include permanently transfected CHO cells, in which agonists and antagonists can be identified by the ability to specifically alter the binding of GTPγ[^{35}S] in membranes prepared from these cells. In another embodiment of the invention, permanently transfected CHO cells could be used for the preparation of membranes which contain significant amounts of the recombinant receptor proteins; these membrane preparations would then be used in receptor binding assays, employing the radiolabelled ligand specific for the particular receptor. Alternatively, a functional assay, such as fluorescent monitoring of ligand-induced changes in internal Ca^{2+} concentration or membrane potential in permanently transfected CHO cells containing each of these receptors individually or in combination would be preferred for HTS. Equally preferred would be an alternative type of mammalian cell, such as HEK293 or COS cells, in similar

formats. More preferred would be permanently transfected insect cell lines, such as *Drosophila* S2 cells. Even more preferred would be recombinant yeast cells expressing the *Drosophila melanogaster* receptors in HTS formats well known to those skilled in the art (*e.g.*, Pausch, *Trends in Biotechnology*, **1997**, *15*, 487-494).

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to nGPCR-x receptors. In one example, the nGPCR-x receptor is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the nGPCR-x receptor and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor compound. In either assay, an inhibitor is identified as a compound that decreases binding between the nGPCR-x receptor and its binding partner. Another contemplated assay involves a variation of the dihybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, some of which are derived from natural products, and some of which arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see Science 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial

libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly encompasses non-peptide modulators, as well as such peptide modulators as neuropeptides other than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified nGPCR-x gene.

The polypeptides of the invention are employed as a research tool for identification, characterization and purification of interacting, regulatory proteins. Appropriate labels are incorporated into the polypeptides of the invention by various methods known in the art and the polypeptides are used to capture interacting molecules. For example, molecules are incubated with the labeled polypeptides, washed to remove unbound polypeptides, and the polypeptide complex is quantified. Data obtained using different concentrations of polypeptide are used to calculate values for the number, affinity, and association of polypeptide with the protein complex.

Labeled polypeptides are also useful as reagents for the purification of molecules with which the polypeptide interacts including, but not limited to, inhibitors. In one embodiment of affinity purification, a polypeptide is covalently coupled to a chromatography column. Cells and their membranes are extracted, and various cellular subcomponents are passed over the column. Molecules bind to the column by virtue of their affinity to the polypeptide. The polypeptide-complex is recovered from the column, dissociated and the recovered molecule is subjected to protein sequencing. This amino acid sequence is then used to identify the captured molecule or to design degenerate oligonucleotides for cloning the corresponding gene from an appropriate cDNA library.

Alternatively, compounds may be identified which exhibit similar properties to the ligand for the nGPCR-x of the invention, but which are smaller and exhibit a

longer half time than the endogenous ligand in a human or animal body. When an organic compound is designed, a molecule according to the invention is used as a "lead" compound. The design of mimetics to known pharmaceutically active compounds is a well-known approach in the development of pharmaceuticals based on such "lead" compounds. Mimetic design, synthesis and testing are generally used to avoid randomly screening a large number of molecules for a target property. Furthermore, structural data deriving from the analysis of the deduced amino acid sequences encoded by the DNAs of the present invention are useful to design new drugs, more specific and therefore with a higher pharmacological potency.

Comparison of the protein sequence of the present invention with the sequences present in all the available databases showed a significant homology with the transmembrane portion of G protein coupled receptors. Accordingly, computer modeling can be used to develop a putative tertiary structure of the proteins of the invention based on the available information of the transmembrane domain of other proteins. Thus, novel ligands based on the predicted structure of nGPCR-x can be designed.

In a particular embodiment, the novel molecules identified by the screening methods according to the invention are low molecular weight organic molecules, in which case a composition or pharmaceutical composition can be prepared thereof for oral intake, such as in tablets. The compositions, or pharmaceutical compositions, comprising the nucleic acid molecules, vectors, polypeptides, antibodies and compounds identified by the screening methods described herein, can be prepared for any route of administration including, but not limited to, oral, intravenous, cutaneous, subcutaneous, nasal, intramuscular or intraperitoneal. The nature of the carrier or other ingredients will depend on the specific route of administration and particular embodiment of the invention to be administered. Examples of techniques and protocols that are useful in this context are, *inter alia*, found in Remington's Pharmaceutical Sciences, 16th edition, Osol, A (ed.), **1980**, which is incorporated herein by reference in its entirety.

The dosage of these low molecular weight compounds will depend on the disease state or condition to be treated and other clinical factors such as weight and condition of the human or animal and the route of administration of the compound. For treating human or animals, between approximately 0.5 mg/kg of body weight to 500 mg/kg of body weight of the compound can be administered. Therapy is

typically administered at lower dosages and is continued until the desired therapeutic outcome is observed.

The present compounds and methods, including nucleic acid molecules, polypeptides, antibodies, compounds identified by the screening methods described herein, have a variety of pharmaceutical applications and may be used, for example, to treat or prevent unregulated cellular growth, such as cancer cell and tumor growth. In a particular embodiment, the present molecules are used in gene therapy. For a review of gene therapy procedures, see *e.g.* Anderson, *Science*, **1992**, *256*, 808-813, which is incorporated herein by reference in its entirety.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing a nGPCR-x natural binding partner associated activity in a mammal comprising administering to said mammal an agonist or antagonist to one of the above disclosed polypeptides in an amount sufficient to effect said agonism or antagonism. One embodiment of the present invention, then, is a method of treating diseases in a mammal with an agonist or antagonist of the protein of the present invention comprises administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize nGPCR-x-associated functions.

In an effort to discover novel treatments for diseases, biomedical researchers and chemists have designed, synthesized, and tested molecules that inhibit the function of protein polypeptides. Some small organic molecules form a class of compounds that modulate the function of protein polypeptides. Examples of molecules that have been reported to inhibit the function of protein kinases include, but are not limited to, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642, published November 26, 1992 by Maguire *et al.*), vinylene-azaindole derivatives (PCT WO 94/14808, published July 7, 1994 by Ballinari *et al.*), 1-cyclopropyl-4-pyridyl-quinolones (U.S. Patent No. 5,330,992), styryl compounds (U.S. Patent No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Patent No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 A1), seleoindoles and selenides (PCT WO 94/03427, published February 17, 1994 by Denny *et al.*), tricyclic polyhydroxylic compounds (PCT WO 92/21660, published December 10, 1992 by Dow), and benzylphosphonic acid compounds (PCT WO 91/15495, published October 17, 1991 by Dow *et al*), all of which are incorporated by reference herein, including any drawings.

Exemplary diseases and conditions amenable to treatment based on the present invention include, but are not limited to, thyroid disorders (*e.g.* thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (*e.g.*, Chron's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (e.g., pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, *etc.*); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (*e.g.*, type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, *etc.*); proliferative diseases and cancers (*e.g.*, different cancers such as breast, colon, lung, *etc.*, and hyperproliferative disorders such as psoriasis, prostate hyperplasia, *etc.*); hormonal disorders (*e.g.*, male/female hormonal replacement, polycystic ovarian syndrome, alopecia, *etc.*); sexual dysfunction, among others.

Compounds that can traverse cell membranes and are resistant to acid hydrolysis are potentially advantageous as therapeutics as they can become highly bioavailable after being administered orally to patients. However, many of these protein inhibitors only weakly inhibit function. In addition, many inhibit a variety of protein kinases and will therefore cause multiple side effects as therapeutics for diseases.

Some indolinone compounds, however, form classes of acid resistant and membrane permeable organic molecules. WO 96/22976 (published August 1, 1996 by Ballinari *et al.*) describes hydrosoluble indolinone compounds that harbor tetralin, naphthalene, quinoline, and indole substituents fused to the oxindole ring. These bicyclic substituents are in turn substituted with polar groups including hydroxylated alkyl, phosphate, and ether substituents. U.S. Patent Application Serial Nos. 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187) and 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon

Docket No. 223/298) and International Patent Publication WO 96/22976, published August 1, 1996 by Ballinari *et al.*, all of which are incorporated herein by reference in their entirety, including any drawings, describe indolinone chemical libraries of indolinone compounds harboring other bicyclic moieties as well as monocyclic moieties fused to the oxindole ring. Applications 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187), 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298), and WO 96/22976, published August 1, 1996 by Ballinari *et al.* teach methods of indolinone synthesis, methods of testing the biological activity of indolinone compounds in cells, and inhibition patterns of indolinone derivatives, both of which are incorporated by reference herein, including any drawings.

Other examples of substances capable of modulating kinase activity include, but are not limited to, tyrphostins, quinazolines, quinoxolines, and quinolines. The quinazolines, tyrphostins, quinolines, and quinoxolines referred to above include well-known compounds such as those described in the literature. For example, representative publications describing quinazolines include Barker *et al.*, EPO Publication No. 0 520 722 A1; Jones et al., U.S. Patent No. 4,447,608; Kabbe *et al.*, U.S. Patent No. 4,757,072; Kaul and Vougioukas, U.S. Patent No. 5, 316,553; Kreighbaum and Comer, U.S. Patent No. 4,343,940; Pegg and Wardleworth, EPO Publication No. 0 562 734 A1; Barker *et al.*, Proc. of Am. Assoc. for Cancer Research 32:327 (1991); Bertino, J.R., Cancer Research 3:293-304 (1979); Bertino, J.R., Cancer Research 9(2 part 1):293-304 (1979); Curtin *et al.*, Br. J. Cancer 53:361-368 (1986); Fernandes *et al.*, Cancer Research 43:1117-1123 (1983); Ferris *et al.* J. Org. Chem. 44(2):173-178; Fry *et al.*, Science 265:1093-1095 (1994); Jackman *et al.*, Cancer Research 51:5579-5586 (1981); Jones *et al.* J. Med. Chem. 29(6):1114-1118; Lee and Skibo, Biochemistry 26(23):7355-7362 (1987); Lemus *et al.*, J. Org. Chem. 54:3511-3518 (1989); Ley and Seng, Synthesis 1975:415-522 (1975); Maxwell *et al.*, Magnetic Resonance in Medicine 17:189-196 (1991); Mini *et al.*, Cancer Research 45:325-330 (1985), Phillips and Castle, J. Heterocyclic Chem. 17(19):1489-1596 (1980); Reece *et al.*, Cancer Research 47(11):2996-2999 (1977); Sculier *et al.*, Cancer Immunol. and Immunother. 23:A65 (1986); Sikora *et al.*, Cancer Letters 23:289-295

(1984); and Sikora *et al.*, Analytical Biochem. 172:344-355 (1988), all of which are incorporated herein by reference in their entirety, including any drawings.

Quinoxaline is described in Kaul and Vougioukas, U.S. Patent No. 5,316,553, incorporated herein by reference in its entirety, including any drawings.

Quinolines are described in Dolle *et al.*, J. Med. Chem. 37:2627-2629 (1994); MaGuire, J. Med. Chem. 37:2129-2131 (1994); Burke *et al.*, J. Med. Chem. 36:425-432 (1993); and Burke *et al.* BioOrganic Med. Chem. Letters 2:1771-1774 (1992), all of which are incorporated by reference in their entirety, including any drawings.

Tyrphostins are described in Allen *et al.*, Clin. Exp. Immunol. 91:141-156 (1993); Anafi *et al.*, Blood 82:12:3524-3529 (1993); Baker *et al.*, J. Cell Sci. 102:543-555 (1992); Bilder *et al.*, Amer. Physiol. Soc. pp. 6363-6143:C721-C730 (1991); Brunton *et al.*, Proceedings of Amer. Assoc. Cancer Rsch. 33:558 (1992); Bryckaert *et al.*, Experimental Cell Research 199:255-261 (1992); Dong *et al.*, J. Leukocyte Biology 53:53-60 (1993); Dong *et al.*, J. Immunol. 151(5):2717-2724 (1993); Gazit *et al.*, J. Med. Chem. 32:2344-2352 (1989); Gazit *et al.*, J. Med. Chem. 36:3556-3564 (1993); Kaur *et al.*, Anti-Cancer Drugs 5:213-222 (1994); King *et al.*, Biochem. J. 275:413-418 (1991); Kuo *et al.*, Cancer Letters 74:197-202 (1993); Levitzki, A., The FASEB J. 6:3275-3282 (1992); Lyall *et al.*, J. Biol. Chem. 264:14503-14509 (1989); Peterson *et al.*, The Prostate 22:335-345 (1993); Pillemer *et al.*, Int. J. Cancer 50:80-85 (1992); Posner *et al.*, Molecular Pharmacology 45:673-683 (1993); Rendu *et al.*, Biol. Pharmacology 44(5):881-888 (1992); Sauro and Thomas, Life Sciences 53:371-376 (1993); Sauro and Thomas, J. Pharm. and Experimental Therapeutics 267(3):119-1125 (1993); Wolbring *et al.*, J. Biol. Chem. 269(36):22470-22472 (1994); and Yoneda *et al.*, Cancer Research 51:4430-4435 (1991); all of which are incorporated herein by reference in their entirety, including any drawings.

Other compounds that could be used as modulators include oxindolinones such as those described in U.S. patent application Serial No. 08/702,232 filed August 23, 1996, incorporated herein by reference in its entirety, including any drawings.

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S. Application Serial No. 08/702,282, filed August 23, 1996 and International patent publication number WO 96/22976, published August 1 1996, both of which are incorporated herein by reference in their entirety, including any drawings, figures or

tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC_{50} as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharm-acokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Toxicity studies can also be carried out by measuring the blood cell composition. For example, toxicity studies can be carried out in a suitable animal model as follows: 1) the compound is administered to mice (an untreated control mouse should also be used); 2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and 3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia, Journal of American Veterinary Medical Assoc., 202:229-249, 1993). Representative animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less

preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness. Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

nGPCR-x mRNA transcripts may found in many tissues, including, but not limited to, brain, peripheral blood lymphocytes, pancreas, ovary, uterus, testis, salivary gland, kidney, adrenal gland, liver, bone marrow, prostate, fetal liver, colon, muscle, and fetal brain, and may be found in many other tissues. Within the brain, nGPCR-x mRNA transcripts may be found in many tissues, including, but not limited to, frontal lobe, hypothalamus, pons, cerebellum, caudate nucleus, and medulla. Tissues and brain regions where specific nGPCR mRNA transcripts are expressed are identified in the Examples, below.

Odd numbered nucleotide sequences ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 will, as detailed above, enable screening the endogenous neurotransmitters/hormones/ligands which activate, agonize, or antagonize nGPCR-x and for compounds with potential utility in treating disorders including, but not limited to, thyroid disorders (*e.g.* thyreotoxicosis, myxoedema); renal failure; inflammatory conditions (*e.g.*, Chron's disease); diseases related to cell differentiation and homeostasis; rheumatoid arthritis; autoimmune disorders; movement disorders; CNS disorders (*e.g.*, pain including migraine; stroke; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, anxiety, generalized anxiety disorder, post-traumatic-stress disorder, depression, bipolar disorder, delirium, dementia, severe mental retardation; dyskinesias, such as Huntington's disease or Tourette's Syndrome; attention disorders including ADD and ADHD, and degenerative disorders such as Parkinson's, Alzheimer's; movement disorders, including ataxias, supranuclear palsy, *etc.*); infections, such as viral infections caused by HIV-1 or HIV-2; metabolic and cardiovascular diseases and disorders (*e.g.*, type 2 diabetes, obesity, anorexia, hypotension, hypertension, thrombosis, myocardial infarction, cardiomyopathies, atherosclerosis, *etc.*); proliferative diseases and cancers (*e.g.*, different cancers such as breast, colon, lung,

etc., and hyperproliferative disorders such as psoriasis, prostate hyperplasia, *etc.*); hormonal disorders (*e.g.*, male/female hormonal replacement, polycystic ovarian syndrome, alopecia, *etc.*); sexual dysfunction, among others.

For example, nGPCR-x may be useful in the treatment of respiratory ailments such as asthma, where T cells are implicated by the disease. Contraction of airway smooth muscle is stimulated by thrombin. Cicala *et al* (1999) Br J Pharmacol 126:478-484. Additionally, in bronchiolitis obliterans, it has been noted that activation of thrombin receptors may be deleterious. Hauck *et al.*(1999) Am J Physiol 277:L22-L29. Furthermore, mast cells have also been shown to have thrombin receptors. Cirino *et al* (1996) J Exp Med 183:821-827. nGPCR-x may also be useful in remodeling of airway structure s in chronic pulmonary inflammation via stimulation of fibroblast procollagen synthesis. See, e.g., Chambers *et al.* (1998) Biochem J 333:121-127; Trejo *et al.* (1996) J Biol Chem 271:21536-21541.

In another example, increased release of sCD40L and expression of CD40L by T cells after activation of thrombin receptors suggests that nGPCR-x may be useful in the treatment of unstable angina due to the role of T cells and inflammation. See Aukrust *et al.* (1999) Circulation 100:614-620.

A further example is the treatment of inflammatory diseases, such as psoriasis, inflammatory bowel disease, multiple sclerosis, rheumatoid arthritis, and thyroiditis. Due to the tissue expression profile of nGPCR-x, inhibition of thrombin receptors may be beneficial for these diseases. See, *e.g.*, Morris *et al.* (1996) Ann Rheum Dis 55:841-843. In addition to T cells, NK cells and monocytes are also critical cell types which contribute to the pathogenesis of these diseases. See, e.g., Naldini & Carney (1996) Cell Immunol 172:35-42; Hoffman & Cooper (1995) Blood Cells Mol Dis 21:156-167; Colotta *et al.* (1994) Am J Pathol 144:975-985.

Expression of nGPCR-x in bone marrow and spleen may suggest that it may play a role in the proliferation of hematopoietic progenitor cells. See DiCuccio *et al.* (1996) Exp Hematol 24:914-918.

As another example, nGPCR-x may be useful in the treatment of acute and/or traumatic brain injury. Astrocytes have been demonstrated to express thrombin receptors. Activation of thrombin receptors may be involved in astrogliosis following brain injury. Therefore, inhibition of receptor activity may be beneficial for limiting neuroinflammation. Scar formation mediated by astrocytes may also be limited by inhibiting thrombin receptors. See, *e.g*, Pindon *et al.* (1998) Eur J Biochem 255:766-

774; Ubl & Reiser. (1997) Glia 21:361-369; Grabham & Cunningham (1995) J Neurochem 64:583-591.

nGPCR-x receptor activation may mediate neuronal and astrocyte apoptosis and prevention of neurite outgrowth. Inhibition would be beneficial in both chronic and acute brain injury. See, *e.g.*, Donovan *et al.* (1997) J Neurosci 17:5316-5326; Turgeon *et al* (1998) J Neurosci 18:6882-6891; Smith-Swintosky *et al.* (1997) J Neurochem 69:1890-1896; Gill *et al.* (1998) Brain Res 797:321-327; Suidan *et al.* (1996) Semin Thromb Hemost 22:125-133.

The attached Sequence Listing contains the sequences of the polynucleotides and polypeptides of the invention and is incorporated herein by reference in its entirety.

As described above and in Example 4 below, the genes encoding nGPCR-1 (nucleic acid sequence SEQ ID NO: 1, SEQ ID NO: 73, amino acid sequence SEQ ID NO: 2, SEQ ID NO:74), nGPCR-9 (nucleic acid sequence SEQ ID NO:9, SEQ ID NO:77, amino acid sequence SEQ ID NO:10, SEQ ID NO:78), nGPCR-11 (nucleic acid sequence SEQ ID NO:11, SEQ ID NO:79, amino acid sequence SEQ ID NO:12, SEQ ID NO:80), nGPCR-16 (nucleic acid sequence SEQ ID NO: 21, SEQ ID NO:81, amino acid sequence SEQ ID NO: 22, SEQ ID NO:82), nGPCR-40 (nucleic acid sequence SEQ ID NO:53, SEQ ID NO:83, amino acid sequence SEQ ID NO:54, SEQ ID NO:84), nGPCR-54 (nucleic acid sequence SEQ ID NO:59, SEQ ID NO:85, amino acid sequence SEQ ID NO:60, SEQ ID NO: 86), nGPCR-56 (nucleic acid sequence SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, amino acid sequence SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90), nGPCR-58 (nucleic acid sequence SEQ ID NO:3, SEQ ID NO:185, amino acid sequence SEQ ID NO:4, SEQ ID NO: 186) have been detected in brain tissue indicating that these n-GPCR-x proteins are neuroreceptors. The identification of modulators such as agonists and antagonists is therefore useful for the identification of compounds useful to treat neurological diseases and disorders. Such neurological diseases and disorders, including but are not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia as well as depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like.

Methods of Screening Human Subjects

Thus in yet another embodiment, the invention provides genetic screening procedures that entail analyzing a person's genome -- in particular their alleles for GPCRs of the invention -- to determine whether the individual possesses a genetic characteristic found in other individuals that are considered to be afflicted with, or at risk for, developing a mental disorder or disease of the brain that is suspected of having a hereditary component. For example, in one embodiment, the invention provides a method for determining a potential for developing a disorder affecting the brain in a human subject comprising the steps of analyzing the coding sequence of one or more GPCR genes from the human subject; and determining development potential for the disorder in said human subject from the analyzing step.

More particularly, the invention provides a method of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor, comprising the steps of: (a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering the amino acid sequence, expression, or biological activity of at least one seven transmembrane receptor that is expressed in the brain, wherein the seven transmembrane receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 74, 186, 78, 80, 82, 84, 86, 90, and 94, or an allelic variant thereof, and wherein the nucleic acid corresponds to the gene encoding the seven transmembrane receptor; and (b) diagnosing the disorder or predisposition from the presence or absence of said mutation, wherein the presence of a mutation altering the amino acid sequence, expression, or biological activity of allele in the nucleic acid correlates with an increased risk of developing the disorder. In preferred variations, the seven transmembrane receptor is nGPCR-40 or nGPCR-54 comprising amino acid sequences set forth in SEQ ID NO: 84 for nGPCR-40 and SEQ ID NO: 86 for nGPCR-54, or an allelic variant thereof, and the disease is schizophrenia.

By "human subject" is meant any human being, human embryo, or human fetus. It will be apparent that methods of the present invention will be of particular interest to individuals that have themselves been diagnosed with a disorder affecting the brain or have relatives that have been diagnosed with a disorder affecting the brain.

By "screening for an increased risk" is meant determination of whether a genetic variation exists in the human subject that correlates with a greater likelihood

of developing a disorder affecting the brain than exists for the human population as a whole, or for a relevant racial or ethnic human sub-population to which the individual belongs. Both positive and negative determinations (i.e., determinations that a genetic predisposition marker is present or is absent) are intended to fall within the scope of screening methods of the invention. In preferred embodiments, the presence of a mutation altering the sequence or expression of at least one nGPCR-40 or nGPCR-54 seven transmembrane receptor allele in the nucleic acid is correlated with an increased risk of developing schizophrenia, whereas the absence of such a mutation is reported as a negative determination.

The "assaying" step of the invention may involve any techniques available for analyzing nucleic acid to determine its characteristics, including but not limited to well-known techniques such as single-strand conformation polymorphism analysis (SSCP) [Orita *et al.*, *Proc Natl. Acad. Sci. USA, 86*: 2766-2770 (1989)]; heteroduplex analysis [White *et al.*, *Genomics, 12*: 301-306 (1992)]; denaturing gradient gel electrophoresis analysis [Fischer *et al.*, *Proc. Natl. Acad. Sci. USA, 80*: 1579-1583 (1983); and Riesner *et al.*, *Electrophoresis, 10*: 377-389 (1989)]; DNA sequencing; RNase cleavage [Myers *et al.*, *Science, 230*: 1242-1246 (1985)]; chemical cleavage of mismatch techniques [Rowley *et al.*, *Genomics, 30*: 574-582 (1995); and Roberts *et al.*, *Nucl. Acids Res.*, *25*: 3377-3378 (1997)]; restriction fragment length polymorphism analysis; single nucleotide primer extension analysis [Shumaker *et al.*, *Hum. Mutat.*, *7*: 346-354 (1996); and Pastinen *et al.*, *Genome Res.*, *7*: 606-614 (1997)]; 5' nuclease assays [Pease *et al.*, *Proc. Natl. Acad. Sci. USA, 91*:5022-5026 (1994)]; DNA Microchip analysis [Ramsay, G., *Nature Biotechnology, 16*: 40-48 (1999); and Chee *et al.*, U.S. Patent No. 5,837,832]; and ligase chain reaction [Whiteley *et al.*, U.S. Patent No. 5,521,065]. [See generally, Schafer and Hawkins, *Nature Biotechnology, 16*: 33-39 (1998).] All of the foregoing documents are hereby incorporated by reference in their entirety.

Thus, in one preferred embodiment involving screening nGPCR-40 or nGPCR-54 sequences, for example, the assaying step comprises at least one procedure selected from the group consisting of: (a) determining a nucleotide sequence of at least one codon of at least one nGPCR-40 or nGPCR-54 allele of the human subject; (b) performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; (c) performing a polynucleotide migration assay to

determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; and (d) performing a restriction endonuclease digestion to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences.

In a highly preferred embodiment, the assaying involves sequencing of nucleic acid to determine nucleotide sequence thereof, using any available sequencing technique. [See, *e.g.*, Sanger *et al.*, *Proc. Natl. Acad. Sci. (USA), 74*: 5463-5467 (1977) (dideoxy chain termination method); Mirzabekov, *TIBTECH, 12*: 27-32 (1994) (sequencing by hybridization); Drmanac *et al.*, *Nature Biotechnology, 16*: 54-58 (1998); U.S. Patent No. 5,202,231; and *Science, 260*: 1649-1652 (1993) (sequencing by hybridization); Kieleczawa *et al.*, *Science, 258*: 1787-1791 (1992) (sequencing by primer walking); (Douglas *et al.*, *Biotechniques, 14*: 824-828 (1993) (Direct sequencing of PCR products); and Akane *et al.*, *Biotechniques 16*: 238-241 (1994); Maxam and Gilbert, *Meth. Enzymol., 65*: 499-560 (1977) (chemical termination sequencing), all incorporated herein by reference.] The analysis may entail sequencing of the entire nGPCR gene genomic DNA sequence, or portions thereof; or sequencing of the entire seven transmembrane receptor coding sequence or portions thereof. In some circumstances, the analysis may involve a determination of whether an individual possesses a particular allelic variant, in which case sequencing of only a small portion of nucleic acid -- enough to determine the sequence of a particular codon characterizing the allelic variant -- is sufficient. This approach is appropriate, for example, when assaying to determine whether one family member inherited the same allelic variant that has been previously characterized for another family member, or, more generally, whether a person's genome contains an allelic variant that has been previously characterized and correlated with a mental disorder having a heritable component.

In another highly preferred embodiment, the assaying step comprises performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences. In a preferred embodiment, the hybridization involves a determination of whether nucleic acid derived from the human subject will hybridize with one or more oligonucleotides, wherein the oligonucleotides have nucleotide sequences that correspond identically to a portion of the GPCR gene sequence taught herein, such as

the nGPCR-40 or nGPCR-54 coding sequence set forth in SEQ ID NOS: 83 for nGPCR-40 or 85 for nGPCR-54, or that correspond identically except for one mismatch. The hybridization conditions are selected to differentiate between perfect sequence complementarity and imperfect matches differing by one or more bases. Such hybridization experiments thereby can provide single nucleotide polymorphism sequence information about the nucleic acid from the human subject, by virtue of knowing the sequences of the oligonucleotides used in the experiments.

Several of the techniques outlined above involve an analysis wherein one performs a polynucleotide migration assay, *e.g.*, on a polyacrylamide electrophoresis gel (or in a capillary electrophoresis system), under denaturing or non-denaturing conditions. Nucleic acid derived from the human subject is subjected to gel electrophoresis, usually adjacent to (or co-loaded with) one or more reference nucleic acids, such as reference GPCR-encoding sequences having a coding sequence identical to all or a portion of SEQ ID NOS: 83 or 85 (or identical except for one known polymorphism). The nucleic acid from the human subject and the reference sequence(s) are subjected to similar chemical or enzymatic treatments and then electrophoresed under conditions whereby the polynucleotides will show a differential migration pattern, unless they contain identical sequences. [See generally Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, New York: John Wiley & Sons, Inc. (1987-1999); and Sambrook *et al.*, (eds.), Molecular Cloning, *A Laboratory Manual*, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press (1989), both incorporated herein by reference in their entirety.]

In the context of assaying, the term "nucleic acid of a human subject" is intended to include nucleic acid obtained directly from the human subject (*e.g.*, DNA or RNA obtained from a biological sample such as a blood, tissue, or other cell or fluid sample); and also nucleic acid derived from nucleic acid obtained directly from the human subject. By way of non-limiting examples, well known procedures exist for creating cDNA that is complementary to RNA derived from a biological sample from a human subject, and for amplifying (*e.g.*, via polymerase chain reaction (PCR)) DNA or RNA derived from a biological sample obtained from a human subject. Any such derived polynucleotide which retains relevant nucleotide sequence information of the human subject's own DNA/RNA is intended to fall within the definition of "nucleic acid of a human subject" for the purposes of the present invention.

In the context of assaying, the term "mutation" includes addition, deletion, and/or substitution of one or more nucleotides in the GPCR gene sequence (*e.g.*, as compared to the seven transmembrane receptor-encoding sequences set forth of SEQ ID NOS: 74, 186, 78, 80, 82, 84, 86, 90, and 94) and other polymorphisms that occur in introns (where introns exist) and that are identifiable via sequencing, restriction fragment length polymorphism, or other techniques. The various activity examples provided herein permit determination of whether a mutation modulates activity of the relevant receptor in the presence or absence of various test substances.

In a related embodiment, the invention provides methods of screening a person's genotype with respect to GPCR's of the invention, and correlating such genotypes with diagnoses for disease or with predisposition for disease (for genetic counseling). For example, the invention provides a method of screening for an nGPCR-40 or nGPCR-54 hereditary schizophrenia genotype in a human patient, comprising the steps of: (a) providing a biological sample comprising nucleic acid from the patient, the nucleic acid including sequences corresponding to said patient's nGPCR-40 or nGPCR-54 alleles; (b) analyzing the nucleic acid for the presence of a mutation or mutations; (c) determining an nGPCR-40 or nGPCR-54 genotype from the analyzing step; and (d) correlating the presence of a mutation in an nGPCR-40 or nGPCR-54 allele with a hereditary schizophrenia genotype. In a preferred embodiment, the biological sample is a cell sample containing human cells that contain genomic DNA of the human subject. The analyzing can be performed analogously to the assaying described in preceding paragraphs. For example, the analyzing comprises sequencing a portion of the nucleic acid (e.g., DNA or RNA), the portion comprising at least one codon of the nGPCR-40 or nGPCR-54 alleles.

Although more time consuming and expensive than methods involving nucleic acid analysis, the invention also may be practiced by assaying protein of a human subject to determine the presence or absence of an amino acid sequence variation in GPCR protein from the human subject. Such protein analyses may be performed, *e.g.*, by fragmenting GPCR protein via chemical or enzymatic methods and sequencing the resultant peptides; or by Western analyses using an antibody having specificity for a particular allelic variant of the GPCR.

The invention also provides materials that are useful for performing methods of the invention. For example, the present invention provides oligonucleotides useful as probes in the many analyzing techniques described above. In general, such

oligonucleotide probes comprise 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides that have a sequence that is identical, or exactly complementary, to a portion of a human GPCR gene sequence taught herein (or allelic variant thereof), or that is identical or exactly complementary except for one nucleotide substitution. In a preferred embodiment, the oligonucleotides have a sequence that corresponds in the foregoing manner to a human GPCR coding sequence taught herein, and in particular, the coding sequences set forth in SEQ ID NO: 83 and 85. In one variation, an oligonucleotide probe of the invention is purified and isolated. In another variation, the oligonucleotide probe is labeled, *e.g.*, with a radioisotope, chromophore, or fluorophore. In yet another variation, the probe is covalently attached to a solid support. [See generally Ausubel *et al.* And Sambrook *et al.*, *supra*.]

In a related embodiment, the invention provides kits comprising reagents that are useful for practicing methods of the invention. For example, the invention provides a kit for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor, comprising, in association: (a) an oligonucleotide useful as a probe for identifying polymorphisms in a human nGPCR-40 or nGPCR-54 seven transmembrane receptor gene, the oligonucleotide comprising 6-50 nucleotides that have a sequence that is identical or exactly complementary to a portion of a human nGPCR-40 or nGPCR-54 gene sequence or nGPCR-40 or nGPCR-54 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution; and (b) a media packaged with the oligonucleotide containing information identifying polymorphisms identifyable with the probe that correlate with schizophrenia or a genetic predisposition therefor. Exemplary information-containing media include printed paper package inserts or packaging labels; and magnetic and optical storage media that are readable by computers or machines used by practitioners who perform genetic screening and counseling services. The practitioner uses the information provided in the media to correlate the results of the analysis with the oligonucleotide with a diagnosis. In a preferred variation, the oligonucleotide is labeled.

In still another embodiment, the invention provides methods of identifying those allelic variants of GPCRs of the invention that correlate with mental disorders. For example, the invention provides a method of identifying a seven transmembrane

allelic variant that correlates with a mental disorder, comprising steps of: (a) providing a biological sample comprising nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny; (b) analyzing the nucleic acid for the presence of a mutation or mutations in at least one seven transmembrane receptor that is expressed in the brain, wherein the at least one seven transmembrane receptor comprises an amino acid sequence selected from the group consisting of SEQ ID NOS: 74, 186, 78, 80, 82, 84, 86, 90, and 94 or an allelic variant thereof, and wherein the nucleic acid includes sequence corresponding to the gene or genes encoding the at least one seven transmembrane receptor; (c) determining a genotype for the patient for the at least one seven transmembrane receptor from said analyzing step; and (d) identifying an allelic variant that correlates with the mental disorder from the determining step. To expedite this process, it may be desirable to perform linkage studies in the patients (and possibly their families) to correlate chromosomal markers with disease states. The chromosomal localization data provided herein facilitates identifying an involved GPCR with a chromosomal marker.

The foregoing method can be performed to correlate GPCR's of the invention to a number of disorders having hereditary components that are causative or that predispose persons to the disorder. For example, in one preferred variation, the disorder is schizophrenia, and the at least one seven transmembrane receptor comprises nGPCR-40 having an amino acid sequence set forth in SEQ ID NO: 84 or an allelic variant thereof.

Also contemplated as part of the invention are polynucleotides that comprise the allelic variant sequences identified by such methods, and polypeptides encoded by the allelic variant sequences, and oligonucleotide and oligopeptide fragments therof that embody the mutations that have been identified. Such materials are useful in *in vitro* cell-free and cell-based assays for identifying lead compounds and therapeutics for treatment of the disorders. For example, the variants are used in activity assays, binding assays, and assays to screen for activity modulators described herein. In one preferred embodiment, the invention provides a purified and isolated polynucleotide comprising a nucleotide sequence encoding a nGPCR-40 or nGPCR-54 receptor allelic variant identified according to the methods described above; and an oligonucleotide that comprises the sequences that differentiate the allelic variant from the nGPCR-40 or nGPCR-54 sequences set forth in SEQ ID NOS: 83 and 88. The

invention also provides a vector comprising the polynucleotide (preferably an expression vector); and a host cell transformed or transfected with the polynucleotide or vector. The invention also provides an isolated cell line that is expressing the allelic variant GPCR polypeptide; purified cell membranes from such cells; purified polypeptide; and synthetic peptides that embody the allelic variation amino acid sequence. In one particular embodiment, the invention provides a purified polynucleotide comprising a nucleotide sequence encoding a nGPCR-40 seven transmembrane receptor protein of a human that is affected with schizophrenia; wherein said polynucleotide hybridizes to the complement of SEQ ID NO: 83 under the following hybridization conditions: (a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaCl, 10% dextran sulfate and (b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS; and wherein the polynucleotide encodes a nGPCR-40 amino acid sequence that differs from SEQ ID NO: 84 by at least one residue.

An examplary assay for using the allelic variants is a method for identifying a modulator of nGPCR-x biological activity, comprising the steps of: (a) contacting a cell expressing the allelic variant in the presence and in the absence of a putative modulator compound; (b) measuring nGPCR-x biological activity in the cell; and (c) identifying a putative modulator compound in view of decreased or increased nGPCR-x biological activity in the presence versus absence of the putative modulator.

Additional features of the invention will be apparent from the following Examples. Examples 1, 2, 4, 11, 12, and 13 are actual, while the remaining Examples are prophetic. Additional features and variations of the invention will be apparent to those skilled in the art from the entirety of this application, including the detailed description, and all such features are intended as aspects of the invention. Likewise, features of the invention described herein can be re-combined into additional embodiments that also are intended as aspects of the invention, irrespective of whether the combination of features is specifically mentioned above as an aspect or embodiment of the invention. Also, only such limitations which are described herein as critical to the invention should be viewed as such; variations of the invention lacking limitations which have not been described herein as critical are intended as aspects of the invention.

EXAMPLES

EXAMPLE 1: IDENTIFICATION OF nGPCR-X

A. Database search

The Celera database was searched using known GPCR receptors as query sequences to find patterns suggestive of novel G protein-coupled receptors. Positive hits were further analyzed with the GCG program BLAST to determine which ones were the most likely candidates to encode G protein-coupled receptors, using the standard (default) alignment produced by BLAST as a guide.

Briefly, the BLAST algorithm, which stands for Basic Local Alignment Search Tool is suitable for determining sequence similarity (Altschul *et al.*, J. Molec. Biol., 1990, 215, 403-410, which is incorporated herein by reference in its entirety). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pair (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, supra). These initial neighborhood word hits act as seeds for initiating searches to find HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Extension for the word hits in each direction are halted when: 1) the cumulative alignment score falls off by the quantity X from its maximum achieved value; 2) the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or 3) the end of either sequence is reached. The Blast algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The Blast program uses as defaults a word length (W) of 11, the BLOSUM62 scoring matrix (see Henikoff *et al.*, Proc. Natl. Acad. Sci. USA, 1992, 89, 10915-10919, which is incorporated herein by reference in its entirety) alignments (B) of 50, expectation (E) of 10, M=5, N=4, and a comparison of both strands.

The BLAST algorithm (Karlin *et al.*, Proc. Natl. Acad. Sci. USA, 1993, 90, 5873-5787, which is incorporated herein by reference in its entirety) and Gapped BLAST perform a statistical analysis of the similarity between two sequences. One measure of similarity provided by the BLAST algorithm is the smallest sum

probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a GPCR gene or cDNA if the smallest sum probability in comparison of the test nucleic acid to a GPCR nucleic acid is less than about 1, preferably less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

Homology searches were performed with the program BLAST version 2.08. A collection of 340 query amino acid sequences derived from GPCR's was used to search the genomic DNA sequence using TBLASTN and alignments with an E-value lower than 0.01 were collected from each BLAST search. The amino acid sequences have been edited to remove regions in the sequence that produce non-significant alignments with proteins that are not related to GPCR's.

Multiple query sequences may have a significant alignment to the same genomic region, although each alignment may not cover exactly the same DNA region. A procedure is used to determine the region of maximum common overlap between the alignments from several query sequences. This region is called the consensus DNA region. The procedure for determining this consensus involves the automatic parsing of the BLAST output files using the program MSPcrunch to produce a tabular report. From this tabular report the start and end of each alignment in the genomic DNA is extracted. This information was used by a PERL script to derive the maximum common overlap. These regions were reported in the form of a unique sequence identifier, a start and the end position in the sequence. The sequences defined by these regions were extracted from the original genomic sequence file using the program fetchdb.

The consensus regions were assembled into a non-redundant set by using the program phrap. After assembly with phrap a set of contigs and singletons was defined as candidate DNA regions coding for nGPCR-x. These sequences were then submitted for further sequence analysis.

Further sequence analysis involved the removal of sequences previously isolated and removal of sequences related to olfactory GPCRs. The transmembrane regions for the sequences that remained were determined using a FORTRAN computer program called "tmtrest.all" [Parodi *et al.*, Comput.Appl.Biosci. 5:527-535(1994)]. Only sequences that contained transmembrane regions in a pattern found in GPCRs were retained.

cDNAs were sequenced directly using an ABI377 fluorescence-based sequencer (Perkin-Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI PRISM™ Ready Dye-Deoxy Terminator kit with Taq FS™ polymerase. Each ABI cycle sequencing reaction contained about 0.5 μg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 minute, followed by 50 cycles using the following parameters: 98°C for 30 seconds, annealing at 50°C for 30 seconds, and extension at 60°C for 4 minutes. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using Centriflex™ gel filtration cartridges (Advanced Genetic Technologies Corp., Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which is then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x *g* for 4 minutes at room temperature. Column-purified samples were dried under vacuum for about 40 minutes and then dissolved in 5 μl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for three minutes and loaded into the gel sample wells for sequence analysis using the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the Sequencer program (Gene Codes, Ann Arbor, MI). Generally, sequence reads of 700 bp were obtained. Potential sequencing errors were minimized by obtaining sequence information from both DNA strands and by re-sequencing difficult areas using primers annealing at different locations until all sequencing ambiguities were removed.

The following Table 5 contains the sequences of the polynucleotides and polypeptides of the invention. Start and stop codons within the polynucleotide sequence are identified by boldface type. The transmembrane domains within the polypeptide sequence are identified by underlining.

Table 5

The following DNA sequence beGPCR-**seq1** <SEQ ID NO. 1> was identified in *H. sapiens*:

GTCTGGGGGTGGGGATGCTGGGACAGGGGTCAATTGCCTGAAGCAAGTGCTCTCATCCCCCTAGCTCCTGC
TGATCTAGTTGGGGCTCCAGAGTGGGGAGGAGAAAGGCACTTTGAAACTTCTCTGCCCTTACCGTCTTAGCC
ATCAAACTCTGAGCTGGAGATAGTGACGATGTGACAGGAACTTTCCCTGGGCCTCTCTGGGCCACAATTCCT
GGCCGAGAGAAAGAGGAGGAATGAGGTGAGCACCTTCTTCACTCCTAGGGCCATGTGGTAGAGCTGCAGTCG
CACCTCCTTCTGCCAATAGGCATAGATGAGTGGGTTGAGCAGGGAGTTGCCCACGCCGAGCAGCCACAGGTA
CCGTTCCAGCACTAGGTAGAGGTGACACTCCTGGCAGGCCACCTGCACAATGCCAGTGATAAGGAAGGGGGT
CCAGGATAGAGCAAAGCTCCCAATGAGAACAGACACAGTACGGAGAGCTTTGAAGTCGCTGGGAGTCCGTGG
GGATCGATAACCTCCAGCCATGGCTCCTGCATGTTCCATCTTTCGAATCTGCTGGCTGTGCATGGAGGCAAT
TTGAGCATGTCGCAGTAGAAGAAGACAAAGAGGAGCATGGCTGGGAAGAAGCCAACGCAGGAGAGGGTCAGC
ACGAAGTGAGGGTGAAATACAGCAAAGAAGCTGCACTGCCCTTTGTAGGCAGTCTGCTGGAACATGGGGATT
CCGAGTGGGAGGAAGCCAATGAGGTAAGACACTAACCACAGCCCGGCAATGCAGGCCCCGGCCACGAACCCA
CTCATGATCTTCAAGTAGCGGAAGGGCTGCTTGATGGCAAGGTACCTGTCAAAGGTGATCAGCATGACCGTG
AGGACAGAGGCAGCTGCGGAGGAAGTGACAAATGCCATCCGCAGGCTGCACAGGGTCTTCTGTGTGGGCCGA
GAAGGGCTGGAGAGCTGGTCTGTGAGTAGGCCAGAGATGGCCACACCAATCAAGGTGTCAGCCACAGCCAGA
TTCAAGGTGAAGCAGAGACTGACACCATCATTCTTGTGGATCAACAGCAGCACAGCCACAGCCACTAGTGTG
TTAGTAGCAATGATGAGGGAGGCCAGGACAGCAAGGATCACTCCAAATGAGAAAGATGATTCCATGTCTCGA
AGTGGCAGGACTTCACTTACCAGGGCATG

The following amino acid sequence <SEQ ID NO. 2> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 1:

MESSFSFGVILAVLASLIIATNTLVAVAVLLLIHKNDGVSLCFTLNLAVADTLIGVAISGLLTDQLSSPSRPT
QKTLCSLRMAFVTSSAAASVLTVMLITFDRYLAIKQPFRYLKIMSGFVAGACIAGLWLVSYLIGFLPLGIPMF
QQTAYKGQCSFFAVFHPHFVLTLSCVGFFPAMLLFVFFYCDMLKIASMHSQQIRKMEHAGAMAGGYRSPRTPS
DFKALRTVSVLIGSFALSWTPFLITGIVQVACQECHLYLVLERYLWLLGVGNSLLNPLIYAYWQKEVRLQLYH
MALGVKKVLTSFLLFLSARNCGPERPRESSCHIVTISSSEFDG

The following DNA sequence beGPCR-**seq3**<SEQ ID NO. 3> was identified in *H. sapiens*:

CAGCGCGAGCGCCTTCATGGTGACGGTGTCCATGCGCTGGCAGTGTCTGCGTGCCACCCGGTGCACCTGGAG
CGAGGTGAGGCAGAGCACCGCCAGCGGCAGCACGAAGCCCACGGCATGGAGCGTGGCGGTGAAGGCTGCGAA
GCGCGGACGCTCAGGCTCGGGCGGCAGGCGCAGCGAACAGGACGCGAAGGCGCTGCTGTAGCCAAGCCACGA
GCAGCCAAGTGCAGCGCCTGAGAAGGCCAGCGACTGTCCCCAGGCACAGCCCAGCAGCAGGCCGGCATAGCG
CGGTCGCAGGCGTCCGGCGTAGCGCAGTGGGAAGCCCACTGCCAGCCACTGGTCTGCGCTCAGCGCCGCCAC
GCTCAGCGCCGCGTTGGACGCCAGGAAGGTGTCCAGGAAGCCAATGACTTGGCATGCGCCGGGCGCCGACGG
TGTCCGCCCGCGCATCACACCGAGCAGCGTGAAGGGCATGTCCAGCGCCGCCAGCAGCAGGTGGCCCAGAGA
CAGATTCACCAGGAGGACGCCTGAGGCTCGAGTGCGGAGCTCAGCGCTGTAGGCGCAACAAAGCAGCACCAG
TGCGTTGGATAGCAGCGCCACGGCCAGTACCATCACCAGGAGACCCGCCAGCAGCGCCTCGCCGGGGCCCAT
GGCGCTAGC

The following amino acid sequence <SEQ ID NO. 4> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 3:

SAMGPGEALLAGLLVMVLAVALLSNALVLLCCAYSAELRTRASGVLLVNLSLGHLLLAALDMPFTLLGVMRGR
TPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRLRPRYAGLLLGCAWGQSLAFSGAALGC
SWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVGFVLPLAVLCLTSLQVHRVARRHCQRMDTVTMKALA

The following DNA sequence beGPCR-**seq4** <SEQ ID NO. 5> was identified in *H. sapiens*:

TGTGCAGGTGTGATCTCCATTCCTTTGTACATCCCTCACACGCTGTTCGATGGGATTTTGGAAAGGAAATCT
GTGTATTTTGGCTCACTACTGACTATCTGTTATGTACAGCATCTGTATATAACATTGTCCTCATCAGCTATG
ATCGATACCTGTCAGTCTCAAATGCTGTAAGTCGAACACATTAATTTATCCCCCTTAGAAGATTATGTAAAT
GTATA

The following amino acid sequence <SEQ ID NO. 6> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 5:

CAGVISIPLYIPHTLFEWDFGKEICVFWLTTDYLLCTASVYNIVLISYDRYLSVSNAVSRTHFIPLR
RLCKCI

The following DNA sequence beGPCR-**seq5** <SEQ ID NO. 7> was identified in *H. sapiens*:

GACGTCGAAGCAGGTGATGATGCCCAGGGCGTGCACCGGGTAGGTGAGATCGGTGCGCGCCAGCGGGGACAGG
GCGGTCAGGAGCAGCAGCCAGGTCCCTGCACACGCGGCCACCGCGTAACGACGGCGGCGCCAGCGCTTGGAGC
TGAGCGGGTACAGGATCCCCAGGAAGCGCTCCACGCTGATACAGGTCATGGTGAGGATGCTGGAATACATGTT
TGCGTAAAAGGCCACGGTCACCACGTTGCAAAGCAGCACCCCGAATACCCAGTGGTGGCGGTTGCAATGGTAG
TAGATTTGGAAAGGCAACACGCTGGCCAGCATCAGGTCCGTGACGCTCAGGTTGATCATGAAGATGACCGACG
GGGATCTGGGCCCCATGCGCCGGCACAGCACCCACAGAGAGAAGAGGTTGCCCGGGATGCTGACCGCCGCCAC
CAGCGAGTACACCACGGGCAGGGCCACCGCGATCGCCGGGTTCCGCAGCATCTGCAGCGTCGCGTTGTC

The following amino acid sequence <SEQ ID NO. 8> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 7:

DNATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRSPSVIFMINLSVTDLMLASVLPFQIYY
HCNRHHWVFGVLCNLVVTVAFYANMYSSILTMTCISVERFLGILYPLSSKRWRRRRYAVAACAGTWLLLLTAL
SPLARTDLTYPVHALGIITCFDV

The following DNA sequence beGPCR-**seq9**<SEQ ID NO. 9> was identified in *H. sapiens*:

CCCATGTTCCTGCTCCTGGGCAGCCTCACGTTGTCGGATCTGCTGGCAGGCGCCGCCTACGCCGCCAACAT
CCTACTGTCGGGGCCGCTCACGCTGAAACTGTCCCCCGCGCTCTGGTTCGCACGGGAGGGAGGCGTCTTCG
TGGCACTCACTGCGTCCGTGCTGAGCCTCCTGGGCATCGCGCTGGAGCGCAGCCTCACCATGGCGCGCAGG
GGGCCCGCGCCCGTCTCCAGTCGGGGGCGCACGCTGGCGATGGCAGCCGCGGCCTGG

The following amino acid sequence <SEQ ID NO. 10> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 9:

PMFLLLGSLTLSDLLAGAAYAANILLSGPLTLKLSPALWFAREGGVFVALTASVLSLLGIALERSLTMARRGP
APVSSRGRTLAMAAAAW

The following DNA sequence beGPCR-**seq11** <SEQ ID NO. 11> was identified in *H. sapiens*:

CTGCTCATTGTGGCCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTCTGCTTCCACAT
GAAGACCTGGAAGCCCAGCACTGTTTACCTTTTCAATTTGGCCGTGGCTGATTTCCTCCTTATGATCTGCC
TGCCTTTTCGGACAGACTATTACCTCAGACGTAGACACTGGGCTTTTGGGGACATTCCCTGCCGAGTGGGG
CTCTTCACGTTGGCCATGAACAGGGCCGGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTT
CAAAGTGGTCCACCCCCACCACGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCC
TGTGGGCCCTGGTCATCCTGGGAACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAGACGGCC
GTCTCCTGTGAGAGCTTCATCATGGAGTCGGCCAATGGCTGGCATGACATCATGTTCCAGCTGGAGTTCTT
TATGCCCCTCGGCATCATCTTATTTTGCTCCTTCAAGATTGTTTGGAGCCTGAGGCGGAGGCAGCAGCTGG
CCAGACAGGCTCGGATGAAGAAGGCGACCCGGTTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTAC
CTGCCCAGCGTGTCTGCTAGACTCTATTTCCTCTGGACGGTGCCCTCGAGTGCCTGCGATCCCTCTGTCCA
TGGGGCCCTGCACATAACCCTCAGCTTCACCTACATGAACAGCATGCTGGATCCCCTGGTGTATTATTTTT
CAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTCAAAATCTGCAGTCTGAAACCCAAGCAGCCAGGACAC
TCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCG

The following amino acid sequence <SEQ ID NO. 12> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 11:

LLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTDYYLRRRHWAFGDIPCRVGLF
TLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILGTVYLLLENHLCVQETAVSCE
SFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKATRFIMVVAIVFITCYLPSVSA

RLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKLKICSLKPKQPGHSKTQRPEE
MPIS

The following DNA sequence beGPCR-**seq12**<SEQ ID NO. 13> was identified in *H. sapiens*:

TGGAGCTGTGCCACCACCTATCTGGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTATTGCCCTTCCT
CATCATCACCTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTCCTGT
TCTATATCAACCTTTACGGCAGCATCCTGCTGCTGACCTGCATCTCTGTGCACCAGTTCCTAGGTGTGTGC
CACCCACTGTGTTCGCTGCCCTACCGGACCCGCAGGCATGCCTGGCTGGGCACCAGCACCACCTGGGCCCT
GGTGGTCCTCCAGCTGCTGCCCACACTGGCCTTCTCCCACACGGACTACATCAATGGCCAGATGATCTGGT
ATGACATGACCAGCCAAGAGAATTTTGATCGGCTTTTTGCCTACGGCATAGTTCTGACATTGTCTGGCTTT
CTTTCCCTCCTTGGTCATTTTGGTGTGCTATTCACTGATGGTCAGGAGCCTGATCAAGCCAGAGGAGAACC
TCATGAGGACAGG

The following amino acid sequence <SEQ ID NO. 14> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 13:

WSCATTYLVNLMVADLLYVLLPFLIITYSLDDRWPFGELLCKLVHFLFYINLYGSILLLTCISVHQFLGVCHP
LCSLPYRTRRHAWLGTSTTWALVVLQLLPTLAFSHTDYINGQMIWYDMTSQENFDRLFAYGIVLTLSGFLSLL
GHFGVLFTDGQEPDQARGEPHEDR

The following DNA sequence beGPCR-**seq14**<SEQ ID NO. 15> was identified in *H. sapiens*:

CCACCACGCGCAGCACGCCGACAGGGCCTCTCCCTCCCATTCTCCCGCAGGCCCGGACGACCACGCTGCCT
CCAGCCGGTCGGCAAACTAGGGCAGCTCGCAGCCCACGAACAGCAGCCCCAGCAGCTGGCTCATCTTCAGG
CTCTGCACCTTGGCGCGGGGCATCGCGCTGGGCGCACGGGCTCCACCTGGGCTCGCCGACCAGGCCGCTGC
ACCCGCTGGGGCCTTCAGCCGGTGCCGCCACCAGACGGAGAGTAGGTGGCCACAAGCGACACCCATGATCT
TAACAGGCGCGACGAAGCCCGCGACGGCCTCATAGAACGCGTACACCTGCACGTGCCAGCGCTGCAGGAGC
GCGAAGATCCAGTGGCAGCGACGCATCCCCGGCCAGGCTCGGGCGGAGAGTGGCGCGCCTGGCTGCAGAGA
CGTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAGTACTAGCGCACCACAAACCCCGACCCCCGCGCCA
GCAGCAGTGCCAGCAGCCAGCCCAGGGCGGCGAGGGCACGCGCGGGCAGCGGCCGGCCGTGCGGAAGACGC
ACCGCGCGCCGGCGCTCGAGGGCGATGAGCACCACGAGGTGGGCCGAGGCGCCCCGCCCGGATGCCTGCAG
CAGCTGCAGGAAGCGGCACGCCAGGTCCCCCGTGGCCGCGCGGGGCTCGCCCAGCAGTTCCCAGGCCAGCT
GTGACAGCGCCGTGCCCCCGCACGCGTACAGGTCCGCCAGGGCCAGCTGCACCAGCAGGAAGTCCATCTTG
CGACGCTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACAGGCGGCACAGCACTGTGGTGTTGCCTGCCAC
CGCCACCACCAGGATGACCCCCAGGAACACCAGGCGGACGCG

The following amino acid sequence <SEQ ID NO. 16> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 15:

RVRLVFLGVILVVAVAGNTTVLCRLXXXXXXXXXXKRRKMDFLLVQLALADLYACGGTALSQLAWELLGEPRA
ATGDLACRFLQLLQASGRGASAHLVVLIALERRAVRLPHGRPLPARALAALGWLLALLLARGSGFVVRYXXX
XXXXXXXXTSLQPGAPLSARAWPGMRRCHWIFALLQRWHVQVYAFYEAVAGFVAPVKIMGVACGHLLSVWWRH
RLKAPAGAAAWSASPGGARAPSAMPRAKVQSLKMSQLLGLLFVGCELPFADRLEAAWSSGPAGEWEGEALSAC
CAWW

The following DNA sequence beGPCR-**seq15**<SEQ ID NO. 17> was identified in *H. sapiens*:

TCTAAGTTTTTCTCTGAACTTTGAGCCTGTGAAAAAAGAAGGGATGCTGCCTCAGGCCACCCCAGCCTAGA
TACTCACTCTGAGTGCCATGAGGTAGTAGAGGACACTGATGACAGTCATGGGGAGGAGGTAGAATAGGAAG
GAGGTGACCTGGATGATGAAATTGTAGATCCACATGGGCTTGATGACCGTACAGGTGGCCGAACCTGGGAC
CAGGGACCCATTGGGGAAGTAGTGGAACTTGATGCCATGGATGCTGGTGTTGGGCAGGGAGAAGAGCACGG
AGAAGCCCCAGACGATGCCGAGGATCCTGAGGGCCCGGCGCCGGGTGCTCTGCAGTTTGGCGCGGAACGGG
TGTAGGATGGCCACGTAGCGCTCCACGCTGACGGTGGTGATGCTGAGGATGGAGGCGAAGCACACGGTCTC
AAAGAGGGCCGTCTTGAAGTAGCAGCCCACGGGCCCGAACAAGAAAGGGTAGTTGCGCCACATCTCATAGA
CCTCCAGGGGCATTCCAAGGAGCAGGACCAGGAGGTCAGAGACCGCCAGGCTGAAGAGGTAGTAGTTGGTG
GGCGTCTTCATAGCCTGGTGCTGCAGAATCACCAGGCACACCAGGACATTGCCAATGACCCCCACCACAAA
AATTGGCACATACACCACAGACACGGGGAGGAAGAAGTGGCTGCGCCGAGGTCCGCAGAGGAAGGCCAGAT

ACTCCTCGGTGCTGTTCAGGTGTTTCTGGAATGGATCTTCTAGTTTCTGCTGGTAGATCCAGGAAGCATTC
TGAAGTTTTTCCATCCCTGA

The following amino acid sequence <SEQ ID NO. 18> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 17:

SGMEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLPVSVVYVPIFVVGVIGNVLVCLVILQHQ
AMKTPNTYYLFSLAVSDLLVLLLGMPLEVYEMWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVA
ILHPFRAKLQSTRRRALRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVT
SFLFYLLPMTVISVLYYLMALRVSIAGVAG

The following DNA sequence beGPCR-**seq18** <SEQ ID NO. 19> was identified in *H. sapiens*:

ATCAAGATGATTTTTGCTATCGTGCAAATTATTGGATTTTCCAACTCCATCTGTAATCCCATTGTCTATGC
ATTTATGAATGAAAACTTCAAAAAAAATGTTTTGTCTGCAGTTTGTTATTGCATAGTAAATAAAACCTTCT
CTCCAGCACAAAGGCATGGAAATTCAGGAATTACAATGATGCGGAAGAAAGCAAAGTTTTCCCTCAGAGAG
AATCCAGTG

The following amino acid sequence <SEQ ID NO. 20> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 19:

IKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYCIVNKTFSPAQRHGNSGITMMRKKAKFSLRENP

The following DNA sequence beGPCR-**seq16** <SEQ ID NO. 21> was identified in *H. sapiens*:

GCCACAGCATGCAGTTTTCTGTAGAATTCCACTTTGTCTTTGCACTTGAAGAAGATGAGGTATCTGGTGAC
CAGGATCACCACATAGAATAGGAACCGTGAGGTACATGTGGATGTGCAGCATGGCACTCACAAATTTGCAG
AAGGGCAGCCCAAACATCCAAGTCTTCTTGATGAGGTAGGTCAAGCGAAATGGCACTGTCAGCAGAAAAAC
GCTGTGGACCACCACCAAGTTAATGACCGCCATGGTGGTCACTGACCGGGTGTTCATTTTCACCAGGAGGA
AAAGAATGGAAATGACACCCACCAGCCCGCCAATAAGCACTATGAAGTAGAGGCTGATTAAGTGGGGTGTC
ACTATAGGATCGCAAGAGGAATTCCTGGAGGTATTGTGGCCAGGCATACTTGGGAAGTCACCTGGAGGAGA
AAAAGCACCAGAGTAACTGAC

The following amino acid sequence <SEQ ID NO. 22> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 21:

VSYSGAFSPPGDFPSMPGHNTSRNSSCDPIVTPHLISLYFIVLIGGLVGVISILFLLVKMNTRSVTTMAVINL
VVVHSVFLLTVPFRLTYLIKKTWMFGLPFCKFVSAMLHIHMYLTVPILCGDPGHQIPHLLQVQRQSGILQKTA
CCG

The following DNA sequence beGPCR-**seq17**<SEQ ID NO. 23> was identified in *H. sapiens*:

ACTGACCAAGGTCAGGGCATCGACTGAGGCTAGAAGGCCACAGGAAATGCCAGTCAAGGTGTTGGCGCCTG
CAATCGCACCTACCACAAACTTGACCGGGGGCAGGGGGGCAGGCCCGCCAGCGAACACGGTCAGCAGCACC
AGTCCATTGCAGAGCACGGAGAGCAACACGATGGCCCACACGGCCAGGCGGATGCCCCAGCTTTCAAAGAG
GTACTCACA

The following amino acid sequence <SEQ ID NO. 24> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 23:

CEYLFESWGIRLAVWAIVLLSVLCNGLVLLTVFAGGPAPLPPVKFVVGAIAGANTLTGISCGLLASVDALTLV
S

The following DNA sequence beGPCR-**seq20** <SEQ ID NO. 25> was identified in *H. sapiens*:

AACCCCATCATCTACACGCTCACCAACCGCGACCTGCGCCACGCGCTCCTGCGCCTGGTCTGCTGCGGACG
CCACTCCTGCGGCAGAGACCCGAGTGGCTCCCAGCAGTCGGCGAGCGCGGCTGAGGCTTCCGGGGGCCTGC

GCCGCTGCCTGCCCCCGGGCCTTGATGGGAGCTTCAGCGGCTCGGAGCGCTCATCGCCCCAGCGCGACGGG
CTGGACACCAGCGGCTCCACAGGCAGCCCCGGT

The following amino acid sequence <SEQ ID NO. 26> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 25:

NPIIYTLTNRDLRHALLRLVCCGRHSCGRDPSGSQQSASAAEASGGLRRCLPPGLDGSFSGSERSSPQRDGLD
TSGSTGSPG

The following DNA sequence beGPCR-**seq21** <SEQ ID NO. 27> was identified in *H. sapiens*:

CGTGAAGAACAGCGCCACCATGACCAGCATGTGCACCACGCGCGCTCTGCGCCGCGATGCTCGCGGGTCCG
CAGCCTCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGGCAGAGCTTGCGCGCGATGCGGGCGTACATGACC
ACGATGAGCGCCAGCGGCGCCAGGTAGATGTGCGAGAAGAGCACAGTGGTGTAGACCCTGCGCATGCCCTT
CTCGGGCCAGGCCTCCCAGCAGGAGTAGAGAGGGTAGGAGCGGTTGCGGGCGTCCACCATGAAGTGGTGCT
CCTCACGGGTGACGGTCAGCGTGACGGCCGAGGGACACATGATGAGCAGCGCCAGGGCCCAGATGACGGCG
ATGGTGACGAGCGCCTTCCGCAGGGTCAGCTTCTCGCGGAAAGGGTGCACGATGCAGCGGAACCT

The following amino acid sequence <SEQ ID NO. 28> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 27:

FRCIVHPFREKLTLRKALVTIAVIWALALLIMCPSAVTLTVTREEHHFMVDARNRSYPLYSCWEAWPEKGM
RRVYTTVLFSHIYLAPLALIVVMYARIARKLCXXXXXXXXXXEAADPRASRRRARVVHMLVMVALFFT

The following DNA sequence beGPCR-**seq22**<SEQ ID NO. 29> was identified in *H. sapiens*:

GCAGGGGGCGTGAGTCCTCAGGCACTTCTTGAGGTCCTTGTTGAGCAGGAAGCAGACAATTGGGTTGACGG
CAGCCTGGGCGAAGCTCATCCAAACAGCATGGCCAGGTAGCGGTGGGGCACAGCACAGGCTTTCACAAACA
CTCGCCAGTAGCAGGCCACGATGTAGGGTGACCAGAGGAGCAGAAAGAGCAGTGTGATCGCGTAGAACATG
CGGCCCAGCTGCTTTTCACCCTTGACCTCGTCCATGCCCAGTAGCCGCCGGCTGGCTGCATGCCCATTCTG
CCGGATACCCAGCAGGGTTGGTGGCATGGGCCC

The following amino acid sequence <SEQ ID NO. 30> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 29:

GPMPPTLLGIRQNGHAASRRLLGMDEVKGEKQLGRMFYAITLLFLLLWSPYIVACYWRVFVKACAVPHRYLAT
AVWMSFAQAAVNPIVCFLLNKDLKKCLRTHAPC

The following DNA sequence beGPCR-**seq24** <SEQ ID NO. 31> was identified in *H. sapiens*:

TATTCTGTAATGAAGAATGTCATTCACACTGCCATTGGCACATCCAGTGGCCTCACCTAGCATTGTGAAAG
CCCTTCGGTTGGTGTATTGCCACTTCATTTTAAAAGGATGCACAAGTCCCTGGTGCCTTTCCACAGCAATG
CAGGTCATAGTGAGGATTTCTGTCACAACAGCGGTAGACTGGACAAATGGCACCATCTTGCAAATGAAAGC
ACCTGCAGTAAGGAAATAGGATAAATCATACATCAAAACAAAAAGAATAAAGGTTTCATCTGTGTCTTTGT
AATTATCACTATCAGTCCATTCTGAGCCTCTGCCAAAAAGTTTGATAATTGTAATTACTCTGTAGACACA

The following amino acid sequence <SEQ ID NO. 32> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 31:

VYRVITIIKLFGRGSEWTDSDNYKDTDETFILFVLMYDLSYFLTAGAFICKMVPFVQSTAVVTEILTMTCIAV
ERHQGLVHPFKMKWQYTNRRAFTMLGEATGCANGSVNDILHYRI

The following DNA sequence beGPCR-**seq27** <SEQ ID NO. 33> was identified in *H. sapiens*:

GAGCAACATGATCTTTTTGAAGTACTTGACGGTGTCGTTCTTGACGGTCACGAAGCACAGAGTGTTGATCA
TGCTGTTGCTCATGGCGATGCACTCGACGATGTAGAAGGCAGTGAGGTAGTGCTTCTCCTTCACAAACACG
GTGGGGAAGAAGTCGCGCACGATGGTGAAGCCGTAGAAGGGCGCCCAGCATAGCACGTAGGCGGTGAGGAT

GCACATGAGCACCAGGACCGTCTTCCTGCGGCAGCGCAGCCTCTTGCGGATCTGCTCTGTCTGGAATCCAG
GGACCGCCTTGAACCAGAGCTCCCGGGAGATCCTGGCATAGCACAGGGTCATGGTGACCACGGGGCCCACG
AATTCTATGCCAAAGATAAAGAGGAAGTAGGACTTGTAGTAGAGCTGCTGGTCCACAGGCCAGATCTGGCC
GCAGAAGATCTTTTCCTGGCTCTTGACAATGACGAGGACCGTCTCGGTGGTGAAGTAGGCGGAAGGGATGG
CGATCAGGATGGACACCGTCCACACCAAGGCAATCAGGCCAGTGGCTGTTTGGCACTTCATTCGTGGTCTC
AGCGGATGGACAATAGCCAGATACCTAGGGCAAGAACACAAGTGGAGGCAGCC

The following amino acid sequence <SEQ ID NO. 34> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 33:

GCLHLCSCPRYLAIVHPLRPRMKCQTATGLIALVWTVSILIAIPSAYFTTETVLVIVKSQEKIFCGQIWPVDQ
QLYYKSYFLFIFGIEFVGPVVTMTLCYARISRELWFKAVPGFQTEQIRKRLRCRRKTVLVLMCILTAYVLCWA
PFYGFTIVRDFFPTVFVKEKHYLTAFYIVECIAMSNSMINTLCFVTVKNDTVKYFKKIMLL

The following DNA sequence beGPCR-**seq28** <SEQ ID NO. 35> was identified in *H. sapiens*:

CAGCCACACTGCAGTGATGAAATCAAATGTCCAACACCAACCATAGTCACCATTACTAACTAAGAAGCCAC
AAAACTTCCCTTCCAGGGTGTTCAGCAGCAGGGACAGGGCCCAGGGCAGGGCACACATGACAGTTGACAGG
TTTCTTGGGCAGCAGCAGCAGTACCAGATAGGCCGCAGGACAGACAGGCAGCACTCAGTACTGATGGCACT
CAGCATGCTCAGGCCTACAAGGTAGGCAAAGGTCATCACGCTGGTGAAGAAGCTAGGGAAATTGATGGAGA
TGGAACAGAAGAAGTTACTGAGGTACACCAGGCAATTTATAATCTGGAAGCAGAGGAAGAGGAAGTCGGCC
CCGGCCAGGCTGAGGACGTAGACAGAGAAGGCGTTCCTGCGCATGCGGAAGCCCAGGAGCCAGAGCACAAA
CCCGTTTCCTACCAGCCCGACCAGGGCAATGAAAAGGATCAGGAAGACCGGGATCAG

The following amino acid sequence <SEQ ID NO. 36> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 35:

LIPVFLILFIALVGLVGNGFVLWLLGFRMRRNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSISINFPS
FFTSVMTFAYLVGLSMLSAISTECCLSVLRPIWYCCCCPRNLSTVMCALPWALSLLLNTLEGKFCGFLVSNGD
YGWCWTFDFITAVWL

The following DNA sequence beGPCR-**seq31**<SEQ ID NO. 37> was identified in *H. sapiens*:

GAGAGTCTGATTCTGACTTACATCACATATGTAGGCCTGGGCATTTCTATTTGCAGCCTGATCCTTTGCTTGT
CCGTTGAGGTCCTAGTCTGGAGCCAAGTGACAAAGACAGAGATCACCTATTTACGCCATGTGTGCATTGTTAA
CATTGCAGCCACTTTGCTGATGGCAGATGTGTGGTTCATTGTGGCTTCCTTTCTTAGTGGCCCAATAACACAC
CACAAGGGATGTGTGGCAGCCACATTTTTTGGTCATTTCTTTTACCTTTCTGTATTTTTCTGGATGCTTGCCA
AGGCACTCCTTATCCTCTATGGAATCATGATTGTTTTC

The following amino acid sequence <SEQ ID NO. 38> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 37:

ESLILTYITYVGLGISICSLILCLSVEVLVWSQVTKTEITYLRHVCIVNIAATLLMADVWFIVASFLSGPITH
HKGCVAATFFGHFFYLSVFFWMLAKALLILYGIMIVF

The following DNA sequence beGPCR-**seq32** <SEQ ID NO. 39> was identified in *H. sapiens*:

TTGTGTGGCAGTAGAGAGATGTCAGGCTTCAGAGTCAACAAGAACTGGATTTCAAACTGGATTTGAGGACCCC
CACCTTTGGTAAGTGACTTATTATCTGCGAGCCTCTGTTTCTCTCTTCTTTAAATGAGGACAGTAAATCCCAT
ACGGCAGGGTGGTGGGGAGAATCAGAGATGATACAGCTGGTGATCACATCTGGTTTGTGTTCCCAGGGGCACC
AGACTAGGGTTTCTGAGCATGGATCCAACCGTCCCAGTCTTCGGTACAAAACTGACACCAATCAACGGACGTG
AGGAGACTCCTTGCTACAATCAGACCCTGAGCTTCACGGTGCTGACGTGCATCATTTCCCTTGTCGGACTGAC
AGGAAACGCGGTAGTGCTCTGGCTCCTGGGCTACCGCATGCGCAGGAACGCTGTCTCCATCTACATCCTCAAC
CTGGCCGCAGCAGACTTCCTCTTCCTCAGCTTCCAGATTATACGTTCGCCATTACGCCTCATCAATATCAGCC
ATCTCATCCGCAAAATCCTCGTTTCTGTGATGACCTTTCCCTACTTTACAGGCCTGAGTATGCTGAGCGCCAT
CAGCACCGAGCGCTGCCTGTCTGTTCTGTGGCCCATCTGGTACC

The following amino acid sequence <SEQ ID NO. 40> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 39:

LCGSREMSGFRVNKNWISNWIGPPPLVSDLLSASLCFSLLMRTVNPIRQGGGENQRYSWSHLVCVPRGTRLGF
LSMDPTVPVFGTKLTPINGREETPCYNQTLSFTVLTCIISLVGLTGNAVVLWLLGYRMRRNAVSIYILNLAAA
DFLFLSFQIIRSPLRLINISHLIRKILVSVMTFPYFTGLSMLSAISTERCLSVLWPIWY

The following DNA sequence beGPCR-**seq33** <SEQ ID NO. 41> was identified in *H. sapiens*:

ACAGAAAGCAAGGCCACCAGGACCTTAGGCATAGTCATGGGAGTGTTTGTGTTGTGCTGGCTGCCCTTCTTTG
TCTTGACGATCACAGATCCTTTCATTAATTTTACAACCCTTGAAGATCTGTACAATGTCTTCCTCTGGCTAGG
CTATTTCAACTCTGCTTTCAATCCCATTTTATATGGCATGCTTTATCCTTGGTTTCGCAAGGCATTGAGGATG
ATTGTCACAGGCATGATCTTCCACCCTGACTCTTCCACCCTAAGCCTGTTTTCTGCCCATGCTTAGGCTGTGT
TCATCATTCAATAGGACTCTTTTCTGG

The following amino acid sequence <SEQ ID NO. 42> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 41:

TESKATRTLGIVMGVFVLCWLPFFVLTITDPFINFTTLEDLYNVFLWLGYFNSAFNPILYGMLYPWFRKALRM
IVTGMIFHPDSSTLSLFSAHAAVFIIQDSF

The following DNA sequence beGPCR-**seq34**<SEQ ID NO. 43> was identified in *H. sapiens*:

TAGGAATCTCAGAGAAGAAAGTAAGGAACCAGAAAACCATAAAAGAATGTAAATGGAAAAGAATCAGCAAATC
TTATTCACTTATCACTAAATCTAAAATATGTCAAAATACATGAAGACAACAAATGCTTTAGAACAACTGTTGA
ATGTATTGTCCTACAACTTGGCATATGATCATGCTTGCCTCTCTATGTCCAAGTGTTTATTTTTGCAGTTGAC
CTTAATTTCAAGTTAGTTTTGAGGTCTCTACAGTAATGTTTTTAATCTGTCTCTACTTCTTCAGAAAATAAAT
TAGTTGTTGACGAATCAGTCCTTAAGACCTTGCCGCTTACAATAAGTTTTATTGCCTTCCCAAACCATTGGTA
AAAGAAAGCATAAATCAAGGGGTTCATAGCTGAATTATAATAAACACACCAAACTAAAATCTCATAAACATAA
GGAGGAGTTATAAAATTCATATAAGCATCAATCACTGCATCAACGAGGTATGGTAGCCAAGAGACAAGAAATG
CTGC

The following amino acid sequence <SEQ ID NO. 44> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 43:

LHQRGMVAKRQEMLAAFLVSWLPYLVDAVIDAYMNFITPPYVYEILVWCVYYNSAMNPLIYAFFYQWFGKAIK
LIVSGKVLRTDSSTTNLFSEEVETDKHYCRDLKTNLKLRSTAKINTWTRGKHDHMPSCRTIHSTVVLKHLLSS
CI

The following DNA sequence beGPCR-**seq35** <SEQ ID NO. 45> was identified in *H. sapiens*:

CTGGAAAGAGGTCCTCGATCTATCCTCTACGCCGTCCTTGGTTTTGGGGCTGTGCTGGCAGCGTTTGGAAACT
TACTGGTCATGATTGCTATCCTTCACTTCTAACAACTGCACACACCTACAAACTTTCTGATTGCGTCGCTGGC
CTGTGCTGACTTCTTGGTGGGAGTCACTGTGATGCCCTTCAGCACAGTGAGGTCTGTGGAGAGCTGTTGGTAC
TTTGGGGACAGTTACTGTAAATTCCATACATGTTTTGACACATCTTTCTGTTTTGCTTCTTTATTTCATTTAT
GCTGTATCTCTGTTGATAGATACATTGCTGTTACTGATCCTCTGACCTATCCAACCAAGTTTACTGTGTCAGT
TTCAGGGATATGCATTGTTCTTTCCTGGTTCTTTTCTGTCACATACAGCTTTTCGATCTTTTACACGGGAGCC
AACGAAGAAGGAATTGAGGAATTAGTAGTTGCTCTAACCTGTGTAGGAGGCTGCCAGGCTCCACTGAATCAAA
ACTGGGTCCTACTTTGTTTTCTTCTATTCTTTATACCCAATGTCGCCATGGTGTTTATATACAGTAAGATATT
TTTGGTGGCCAAGCATCAGGCTAGGAAGATAGAAAGTACAGCCAGCCAAGCTCAGTCCTTCTCAGAGAGTTAC
AAGGAAAGAGTAGCAAAAAGAGAGAGAAAGGCTGCCAAAACCTTGGGAATTGCTATGGCAGCATTTCTT

The following amino acid sequence <SEQ ID NO. 46> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 45:

LERGPRSILYAVLGFGAVLAAFGNLLVMIAILHFQLHTPTNFLIASLACADFLVGVTVMPFSTVRSVESCWYF
GDSYCKFHTCFDTSFCFASLFHLCCISVDRYIAVTDPLTYPTKFTVSVSGICIVLSWFFSVTYSFSIFYTGAN
EEGIEELVVALTCVGGCQAPLNQNWVLLCFLLFFIPNVAMVFIYSKIFLVAKHQARKIESTASQAQSFSESYK
ERVAKRERKAAKTLGIAMAAFL

The following DNA sequence beGPCR-**seq36** <SEQ ID NO. 47> was identified in *H. sapiens*:

AACCAGGTGGCCTTACTCCTAAGACCCCTGGCCTTGTCTATGGCCTTTATCAACAGCTGTCTCAATCCAGTTC
TCTATGTCTTCATTGGGCATGACTTCTGGGAGCACTTGCTCCACTCCCTGCTAGCTGCCTTAGAACGGGCACT
TAGCGAGGAGCCAGATAGTGCCTGAATCCCAGCTCCCAGGCAGATGAGTCCTTTATAACATGACCCAATTTCC
TACTCCATTTTCCCACCACTCAATCCTCTTCCCAAACAGCTCTACCATAATCCAACATCCAACAGAATTTAAG
AGAATAAACCACAACTTTTAAGTGAGCTCTATGTGCTAGGTCATGTTTTAGAATACAACCTTAAGTGCCTGGA
AGATGGAGGCAAGAAACAAACAAGGTCTCATTCTTTAGAGGAAGACAGTTCACCAAGACTCAAACAGAAAAAA
AGATAGTTATCTTGTGACAAAACAAGTCATAAAATTGGGTCAGGACCTGCAGCAATGACTTTATGCTAGAATC
CAGAGCACTAGCAGGAAACTGCTTAAATTTTACTTAATCAAAGTCAAGTTTGGACATACATGTCAGGTAAAAC
CTAGCAGAGATGAGCTACCTTGATTTTAAAACTTCAAGGGATAGCTCAATGTCATCAAGATCCTTTTGATGAC
TTG

The following amino acid sequence <SEQ ID NO. 48> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 47:

NQVALLLRPLALSMAFINSCLNPVLYVFIGHDFWEHLLHSLLAALERALSEEPDSAIPAPRQMSPLHDPISYS
IFPPLNPLPKQLYHNPTSNRIENKPQLLSELYVLGHVLEYNLKCLEDGGKKQTRSHSLEEDSSPRLKQKKRLS
CDKTSHKIGSGPAAMTLCNPEHQETAILLNQSQVWTYMSGKTQRATLILKLQGIAQCHQDPFDDL

The following DNA sequence beGPCR-**seq37**<SEQ ID NO. 49> was identified in *H. sapiens*:

GCTTGTTCACGGCCACCATCCTCAAGCTGTTGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGGAGGCGCGC
GGTGGGCCTGGCCGCGGTGGTCTTGCTGGCCTTTGTCACCTGCTTCGCCCCCAACAACTTCGTGCTCCTGGCG
CACATCGTGAGCCGCCTGTTCTACGGCAAGAGCTACTACCACGTGTACAAGCTCACGCTGTGTCTCAGCTGCC
TCAACAACTGTCTGGACCCGTTTGTTTATTACTTTGCGTCCCGGGAATTCCAGCTGCGCCTGCGGGAATATTT
GGGCTGCCGCCGGGTGCCCAGAGACACCCTGGACACGCGCCGCGAGAGCCTCTTCTCCGCCAGGACCACGTCC
GTGCGCTCCGAGGCCGGTGCGCACCCTGAAGGGATGGAGGGAGCCACCAGGCCCGGCCTCCAGAGGCAGGAGA
GTGTGTTCTGAGTCCCGGGGGCGCAGC

The following amino acid sequence <SEQ ID NO. 50> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 49:

LFTATILKLLRTEEAHGREQRRRAVGLAAVVLLAFVTCFAPNNFVLLAHIVSRLFYGKSYYHVYKLTLCLSCL
NNCLDPFVYYFASREFQLRLREYLGCRRVPRDTLDTRRESLFSARTTSVRSEAGAHPEGMEGATRPGLQRQES
VFVPGAQAAPPGLR

The following DNA sequence beGPCR-**seq38** <SEQ ID NO. 51> was identified in *H. sapiens*:

TTACTTATTCTGCCCTTTATCCAACTTTTAATTCCCTTTGCTATTCTCCTGCCTCATTTTCTGGCCTCATTTT
CCCTATTATCCTGCCTCACATTGATCAAGGGATGAGGCTGGCAGGATCCGGAACCCACAGGGCCCCGTGGGCC
ATGAGAGGCTCCTGGACTTGAACCTCAGGACACTCCCACTCTGGCTGCCGGCAGGGATGGAAGCTGGATGAGC
AGGCAGGAGCTGGCAGTGGGGGTGGAGAGCCATAGGCTATTGGGGTGGACAGGCTTGGGTGCCTCATGGGAGC
TCCCCATGGGAGCTGTGGCCCCTTGGGGCCTCTTATTTCTCACCCCAGGCTTTCCCGGGAGAGGTTCAAGTCA
GAAGATGCCCCAAAGATCCACGTGGCCCTGGGTGGCAGCCTGTTCCTCCTGAATCTGGCCTTCTTGGTCAATG
TGGGGAGTGGCTCAAAGGGGTCTGATGCTGCCTGCTGGGCCCGGGGGGCTGTCTTCCACTACTTCCTGCTCTG
TGCCTTCACCTGGATGGGCCTTGAAGCCTTCCACCTCTACCTGCTCGCTGTCAGGGTCTTCAACACCTACTTC
GGGCACTACTTCCTGAAGC

The following amino acid sequence <SEQ ID NO. 52> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 51:

ETYSALYPTFNSLCYSPASFSGLIFPIILPHIDQGMRLAGSGTHRAPWAMRGSWTTSGHSHSGCRQGWKLDEQ
AGAGSGGGEPAIGVDRLGCLMGAPHGSCGPLGPLISHPRLSRERFKSEDAPKIHVALGGSLFLLNLAFLVNVG
SGSKGSDAACWARGAVFHYFLLCAFTWMGLEAFHLYLLAVRVFNTYFGHYFL

The following DNA sequence beGPCR-**seq40** <SEQ ID NO. 53> was identified in

H. sapiens:

AATTGGTCGGAGAGTGCAGCTGCTTGAAATGGAGGATTGAAATCATCACCAGGAGGTTTCCAAACACAGCCAG
CACAGCCCCAAAGCCAAACACTATGTACAGAATCACCCGGGATCCCGGCGAGAAGGGGATTTTCACACAGGAC
CCATTCACGTTCGCGTAGCACAGCTGCACAGCCACCAGCAGGGATGAATTGCTGCTCATAACGCTGGTATTTA
CATATGGAGAAATTTTGTCCTTGTTGATTATCACAAAAAATACAGGATTGTTCCTGATTTTCATTGCTCCTGC
GGAAAAAAACACATATTCACCAGGATGCCAGAGGAAATGATCA

The following amino acid sequence <SEQ ID NO. 54> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 53:

DHFLWHPGEYVFFSAGAMKIRNNPVFFVIINKDKISPYVNTSVMSSNSSLLVAVQLCYANVNGSCVKIPFSPG
SRVILYIVFGFGAVLAVFGNLLVMISILHFKQLHSPTN

The following DNA sequence beGPCR-**seq41** <SEQ ID NO. 55> was identified in *H. sapiens*:

CACATCTTAACAAGACTGAAAAACATTGATTTGTTTTTAATTTGAAGAGCAATTTATTTGCTATTCATTCATA
GTCTTACTTGATTTTTAAAAACTCATTTCGCTTGGTAATTTTAAAGGTATCCTGAACTTCGTCTATCCAACTG
CTTATATATGTTCAGAAAACAAATTCATGGTTGCTGAACTGTTCTTTAAAACCTGACCAGTTACAATAACTTT
TATTGCTTTCCTAAACCATGGGTAAAATAAAGCATAAATCAAAGGATTCATGGCTGAGTTATAATAAGCACAC
CAACAGCATCATAAATACAGGCAGGGGTTATAAAGCCCATAAAGGCATCAATTAATGAATCAATGCTATATGG
TAACCATGAAATCATAAATGCTACCACTGTGACCCCCAGGGTTTTAGCTGCTTTTCTCTCTCTCCTGGCCACT
CTGGCTTTGTAACTCTCTGAGGATGATTCTGTCTTGCTACCAGTATTTTCTATCTTTTTCGCCTGTCGTCTAG
CCACAAGAAATATGTTACCATACAGAATTATCATAATAAAGGTAGGTATAAAGAAGGATAGAAAATCTGTCAA
CA

The following amino acid sequence <SEQ ID NO. 56> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 55:

LTDFLSFFIPTFIMIILYGNIFLVARRQAKKIENTGSKTESSSESYKARVARRERKAAKTLGVTVVAFMISWL
PYSIDSLIDAFMGFITPACIYEICCWCAYYNSAMNPLIYALFYPWFRKAIKVIVTGQVLKNSSATMNLFSEHI
AVGTKFRIPLKLPSEMSFKSSKTMNEQINCSSNKQINVFQSCDV

The following DNA sequence nGPCR-**seq53** <SEQ ID NO. 57> was identified in *H. sapiens*:

TTTGTGGCAAGGAGACCCTGATCCCGGTCTTCCTGATCCTTTTCATTGCCCTGGTCGGGCTGGTAGGAAACGG
GTTTGTGCTCTGGCTCCTGGGCTTCCGCATGCGCAGGAACGCCTTCTCTGTCTACGTCCTCAGCCTGGCCGGG
GCCGACTTCCTCTTCCTCTGCTTCCAGATTATAAATTGCCTGGTGTACCTCAGTAACTTCTTCTGTTCCATCT
CCATCAATTTCCCTAGCTTCTTCACCACTGTGATGACCTGTGCCTACCTTGCAGGCCTGAGCATGCTGAGCAC
CGTCAGCACCGAGCGCTGCCTGTCCGTCCTGTGGCCCATCTGGTATCGCTGCCGCCGCCCCAGACACCTGTCA
GCGGTCGTGTGTGTCCTGCTCTGGGCCCTGTCCCTACTGCTGAGCATCTTGGAAGGGAAGTTCTGTGGCTTCT
TATTTAGTGATGGTGACTCTGGTTGGTGTCAGACATTTGATTTCATCACTGCAGCGTGGCTGATTTTTTTATT
CATGGTTCTCTGTGGGTCCAGTCTGGCCCTGCTGGTCAGGATCCTCTGTGGCTCCAGGGGTCTGCCACTGACC
AGGCTGTACCTGACCATCCTGCTCACAGTGCTGGTGTCCCTCCTCTGCGGCCTGCCCTTTGGCATTCAGTGGT
TCCTAATATTATGGATCTGGAAGGATTCTGATGTCTTATTTTGTCATATTCATCCAGTTTCAGTTGTCCTGTC
ATCTCTTAACAGCAGTGCCAACCCCATCATTTACTTCTTCGTGGGCTCTTTTAGGAAGCAGTGGCGGSTGCAG
CACCCGATCCTCAAGCTGGCTCTCCAGAGGGCTCTGCAGGACATTGCTGAGGTGGATCACAGTGAAGGATGCT
TCCGTCAGGGCACCCGGAGATTCAAAGAAGCATTCTGGTGTAGGGATGGACCCCTCTACTTCCATCATATATA
TGTGGCTTTGAGAGGCAACTTTGCCCC

The following amino acid sequence <SEQ ID NO. 58> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 57:

CGKETLIPVFLILFIALVGLVGNGFVLWLLGFRMRRNAFSVYVLSLAGADFLFLCFQIINCLVYLSNFFCSIS
INFPSFFTTVMTCAYLAGLSMLSTVSTERCLSVLWPIWYRCRRPRHLSAVVCVLLWALSLLLSILEGKFCGFL
FSDGDSGWCQTFDFITAAWLIFLFMVLCGSSLALLVRILCGSRGLPLTRLYLTILLTVLVSLLCGLPFGIQWF
LILWIWKDSDVLFCHIHPVSVVLSSLNSSANPIIYFFVGSFRKQWRXQHPILKLALQRALQDIAEVDHSEGCF
RQGTRRFKEAFWCRDGPLYFHHIYVALRGNFA

The following DNA sequence nGPCR-**seq54**<SEQ ID NO. 59> was identified in *H. sapiens*:

CTTTGCATCTCACTGTTGAGCAGACAGCCTGCTGAAAGTTGTCGCTGACCACCACATATAGTAACAGGTTACC
AAAGGTGTTCAGAGCAGCATAATGGTCTAGAAACGATGTAAGCTTCATGGATCTGATTCTCAATGGAACAACT
GATTGAAAGCAGGCTGAGATTCGATCCTGAATGACCCTCAAGATATGGAAGGGTAAAAAACATACGTAAAATG
CAAGGAGTAGCAGAATGGTTAGCCTTCGTGCTTTCTGCTTAAGGCAGCTGTCAGTTTGCAGTCCATGGGTCAA
AGTGTGGATAATCGTGGTATAGCAAAGTGTCACTATCACCAAGGGGAGGCAGAAAGTACTTGCAGTCAAAATC
AGGTTGTACCACTTAATAGTATTGAGTTCATCCGAACTGGTGAGGTCGAGACAGGCTGATCTGTTGGTCCTGT
TGGTTGATGTGATCAAGAAGGTCATCGGAATGACAGCTACCAGTGAAATGATCCACACCACAGCACAGGCTAC
AACTGCACATCGAGTTTTGTGAATGGAAAAGCAGCTCATTGGGTGAATGATCACACAGTAGCGGAAG

The following amino acid sequence <SEQ ID NO. 60> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 59:

FRYCVIIHPMSCFSIHKTRCAVVACAVVWIISLVAVIPMTFLITSTNRTNRSACLDLTSSDELNTIKWYNLIL
TASTFCLPLVIVTLCYTTIIHTLTHGLQTDSCLKQKARRLTILLLLAFYVCFLPFHILRVIQDRISACFQSVV
PLRIRSMKLTSFLDHYAALNTFGNLLLYVVVSDNFQQAVCSTVRCK

The following DNA sequence nGPCR-**seq55** <SEQ ID NO. 61> was identified in *H. sapiens*, where the underlined ATG identifies a probable start codon:

GGGAGGGCTCGTAGACACACTAACCCTACCCTTTCTGTTTCTTCCTCATCTTTCCTTTCCATCTGTTTCTCAT
GGTCTCCTGTCTGTCTCTCTCTCTCTCCCCTCTTTCTCTCTCCTCGCTCTTTCTCATCCCCTCCATTTCTGTG
TCAATCTCAATCCATTTATATCGGTGGCCACTTTTCTATCTCTTTGTTCTATCTCTCTCTCTCTCTCTTTCCC
ACTTTGTCTCTGCACGCCTGTTGTGTTTTTCTGCCTGTCTCTCTCTTGCCCTCATCTCTCTGTCTCTCTCTTG
CCCTCATCTCTCTGTCTCTCTGTGTCTGTGTCTCCCCCGCTCATTCCCATTTGCAGGTGCAATGTAGCAGGAC
AACTCATGGAGCCCCCCCGGGCCCATCGAGTACCGGACTGGCTGACCCCCTAGGGTTGGCAGTAGCCCCTGAC
CCTCAGT**ATG**GCCAACACTACCGGAGAGCCTGAGGAGGTGAGCGGCGCTCTGTCCCCACCGTCCGCATCAGCT
TATGTGAAGCTGGTACTGCTGGGACTGATTATGTGCGTGAGCCTGGCGGGTAACGCCATCTTGTCCCTGCTGG
TGCTCAAGGAGCGGGCCCTGCACAAGGCTCCTTACTACTTCCTGCTGGACCTGTGCCTGGCCGATGGCATACG
CTCTGCCGTCTGCTTCCCCTTTGTGCTGGCTTCTGTGCGCCACGGCTCTTCATGGACCTTCAGTGCACTCAGC
TGCAAGATTGTGGCCTTTATGGCCGTGCTCTTTTGCTTCCATGCGGCCTTCATGCTGTTCTGCATCAGCGTCA
CCCGCTACATGGCCATCGCCCACCACCGCTTCTACGCCAAGCGCATGACACTCTGGACATGCGCGGCTG

The following amino acid sequence <SEQ ID NO. 62> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 61:

MANTTGEPEEVSGALSPPSASAYVKLVLLGLIMCVSLAGNAILSLLVLKERALHKAPYYFLLDLCLADGIRSA
VCFPFVLASVRHGSSWTFSALSCKIVAFMAVLFCFHAAFMLFCISVTRYMAIAHHRFYAKRMTLWTCAAE

The following DNA sequence nGPCR-**seq56** <SEQ ID NO. 63> was identified in *H. sapiens*:

AAAAATTGCTGTACTGAACTATTGAATGGAACTTGGAAATAAAGTCCCTTCCAAAATAACTATTCTTCAACAG
AGAGTAATAGGTAAATGTTTTAGAAGTGAGAGGACTCAAATTGCCAATGATTTACTCTTTTATTTTTCCTCCT
AGGTTTCTGGGATAAGTATGTGCAAATAAAAAATAAACATGAGAAGGAACTGTAACCTGATTATGGATTTGGG
AAAAAGATAAATCAACACACAAAGGGAAAAGTAAACTGATTGACAGCCCTCAGGAATGATGCCCTTTTGCCAC
AATATAATTAATATTTCCTGTGTGAAAAACAACTGGTCAAATGATGTCCGTGCTTCCCTGTACAGTTTAATGG
TGCTCATAATTCTGACCACACTCGTTGGCAATCTGATAGTTATTGTTTCTATATCACACTTCAAACAACTTCA
TACCCCAACAAATTGGCTCATTCATTCCATGGCCACTGTGGACTTTCTTCTGGGGTGTCTGGTCATGCCTTAC
AGTATGGTGAGATCTGCTGAGCACTGTTGGTATTTTGGAGAAGTCTTCTGTAAAATTCACACAAGCACCGACA
TTATGCTGAGCTCAGCCTCCATTTTCCATTTGTCTTTCATCTCCATTGACCGCTACTATGCTGTGTGTGATCC
ACTGAGATATAAAGCCAAGATGAATATCTTGGTTATTTGTGTGATGATCTTCATTAGTTGGAGTGTCCCTGCT
GTTTTTGCATTTGGAATGATCTTTCTGGAGCTAAACTTCAAAGGCGCTGAAGAGATATATTACAAACATGTTC
ACTGCAGAGGAGGTTGCTCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCTTTTTATAT
ACCTGGATCTATTATGTTATGTGTCTATTACAGAATATATCTTATCGCTAAAGAACAGGCAAGATTAATTAGT
GATGCCAATCAGA

The following amino acid sequence <SEQ ID NO. 64> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 63:

REKTDQPSGMMPFCHNIINISCVKNNWSNDVRASLYSLMVLIILTTLVGNLIVIVSISHFKQLHTPTNWLIHS
MATVDFLLGCLVMPYSMVRSAEHCWYFGEVFCKIHTSTDIMLSSASIFHLSFISIDRYYAVCDPLRYKAKMNI
LVICVMIFISWSVPAVFAFGMIFLELNFKGAEEIYYKHVHCRGGCSVFFSKISGVLTFMTSFYIPGSIMLCVY
YRIYLIAKEQARLISDANQ

The following DNA sequence nGPCR-**seq57**<SEQ ID NO. 65> was identified in *H. sapiens*:

AACAGTCCCGGGTGGAACCTGGGCATGTATATTTTGATTGTTTTATGCATACTCCTAGTGAAGAACCAATGTC
TTGCTCAGATAGAAGCAAGATACTCAGACTTAGTTTCTCTGTAGCTCCTGCTTTTTATTATTCCTGGTTGGAT
TGCACCACTACTCAGTTTCTATTTTATAATACTGATTATAAAACATGGAGGGAAATAACTTTGTATTGGTTT
TTATGGATAATTTATTATGTGTCCTAGACTCTGGCCTTGTCAAAAGAAGGACGTAAGAAGGCACGATGTATTA
TACTTGGGAATGATAGAAGAGACTGACCTGGTATTTCCACCCGGAAGAGGGAAAGGATTTTAACTACAAATAC
AGGAATCCAGCAGATGGCATCAGAGAACACTATAAAAAAGAAACGATTTGCAACAGCCACCTCTCTTCCAAAA
CAATTCCTTACTTCTGTGGTCTGCAAGGCGGTTTTTTGAATGGAACAGAACATAGTAATATAGGAAAACACAA
TGATGAGAAAAGCCAGCAAGTTCACACCTGTTGGGGAAAAGCACACTTTTAACATCTCAGGCGTAAAAGTCAA
CAGTAAAATTACTGTGGTACAGGTTGAGTATCCCTTACCCAAAATGTTTGAAACCAGAAATGTTTTGGATTTC
GGATTTCGGAATATTTACACATTCATAATGATATATCTTGGAAATGGTTCCCAAGTCTAAACACAAAATTTAT
TTATGTTTCATATACACCTTATACACATAGTCTGAAAGTAATTTTGTACAATATTTTAAATAATTTTGGGCAT
GAAACAAAGTTTGCATACATTGAACCATCAGACAGCAAAAGCTTCAGGTGTGGAATTTTCCACTTGTGGCATC
ATGTTGATGCTCAAAAAGTTCCATATTTTAGAGCATTTCAAATTTTGGATTTTCAAATTACAAATGCTTAACC
TGTACTTAGATGTTAAATACAGTGCCTCTTCCACGGGCACTTTCAGGAAGCATTCTTTTATATAAGCCC

The following amino acid sequence <SEQ ID NO. 66> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 65:

YIKECFLKVPVEEALYLTSKYRLSICNLKIQNLKCSKIWNFLSINMMPQVENSTPEAFAVWFNVCKLCFMPKI
INIVQNYFQTMCIRCININKFCVTWEPFPRYIIMNVIFRNPKSKTFLVSNILGKGYSTCTTVILLLTFTPEML
KVCFSPTGVNLLAFLIIVFSYITMFCSIQKTALQTTEVRNCFGREVAVANRFFFIVFSDAICWIPVFVVKILS
LFRVEIPGQSLLSFPSIIHRAFLRPSFDKARVDTIIHKNQYKVISLPCFIISIIKKLSSGAIQPGIIKSRSYR
ETKSEYLASIARHWFFTRSMHKTIKIYMPRFHPGL

The following DNA sequence nGPCR-**seq58** <SEQ ID NO. 67> was identified in *H. sapiens*:

ACTACCATGGAAGCTGACCTGGGTGCCACTGGCCACAGGCCCCGCACAGAGCTTGATGATGAGGACTCCTACC
CCCAAGGTGGCTGGGACACGGTCTTCCTGGTGGCCCTGCTGCTCCTTGGGCTGCCAGCCAATGGGTTGATGGC
GTGGCTGGCCGGCTCCCAGGCCCGGCATGGAGCTGGCACGCGTCTGGCGCTGCTCCTGCTCAGCCTGGCCCTC
TCTGACTTCTTGTTCCTGGCAGCAGCGGCCTTCCAGATCCTAGAGATCCGGCATGGGGGACACTGGCCGCTGG
GGACAGCTGCCTGCCGCTTCTACTACTTCCTATGGGGCGTGTCCTACTCCTCCGGCCTCTTCCTGCTGGCCGC
CCTCAGCCTCGACCGCTGCCTGCTGGCGCTGTGCCCACACTGGTACCCTGGGCACCGCCCAGTCCGCCTGCCC
CTCTGGGTCTGCGCCGGTGTCTGGGTGCTGGCCACACTCTTCAGCGTGCCCTGGCTGGTCTTCCCCGAGGCTG
CCGTCTGGTGGTACGACCTGGTCATCTGCCTGGACTTCTGGGACAGCGAGGAGCTGTCGCTGAGGATGCTGGA
GGTCCTGGGGGGCTTCCTGCCTTTCCTCCTGCTGCTCGTCTGCCACGTGCTCACCCAGGCCACAGCCTGTCGC
ACCTGCCACCGCCAACAGCAGCCCGCAGCCTGCCGGGGCTTCGCCCGTGTGGCCAGGACCATTCTGTCAGCCT
ATGTGGTCCTGAGGCTGCCCTACCAGCTGGCCCAGCTGCTCTACCTGGCCTTCCTGTGGGACGTCTACTCTGG
CTACCTGCTCTGGGAGGCCCTGGTCTACTCCGACTACCTGATCCTACTCAACAGCTGCCTCAGCCCCTTCCTC
TGCCTCATGGCCAGTGCCGACCTCCGGACCCTGCTGCGCTCCGTGCTCTCGTCCTTCGCGGCAGCTCTCTGCG
AGGAGCGGCCGGGCAGCTTCACGCCCACTGAGCCACAGACCCAGCTAGATTCTGAGGGTCCAACTCTGCCAGA
GCCGATGGCAGAGGCCCAGTCACAGATGGATCCTGTGGCCCAGCCTCAGGTGAACCCCACACTCCAGCCACGA
TCGGATCCCACAGCTCAGCCACAGCTGAACCCTACGGCCCAGCCACAGTCGGATCCCACAGCCCAGCCACAGC
TGAACCTCATGGCCCAGCCACAGTCAGATTCTGTGGCCCAGCCACAGGCAGACACTAACGTCCAGACCCCTGC
ACCTGCTGCC

The following amino acid sequence <SEQ ID NO. 68> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 67:

TTMEADLGATGHRPRTELDDEDSYPQGGWDTVFLVALLLLGLPANGLMAWLAGSQARHGAGTRLALLL
LSLALSDFLFLAAAAFQILEIRHGGHWPLGTAACRFYYFLWGVSYSSGLFLLAALSLDRCLLALCPHW

YPGHRPVRLPLWVCAGVWVLATLFSVPWLVFPEAAVWWYDLVICLDFWDSEELSLRMLEVLGGFLPFL
LLLVCHVLTQATACRTCHRQQQPAACRGFARVARTILSAYVVLRLPYQLAQLLYLAFLWDVYSGYLLW
EALVYSDYLILLNSCLSPFLCLMASADLRTLLRSVLSSFAAALCEERPGSFTPTEPQTQLDSEGPTLP
EPMAEAQSQMDPVAQPQVNPTLQPRSDPTAQPQLNPTAQPQSDPTAQPQLNLMAQPQSDSVAQPQADT
NVQTPAPAA

The following DNA sequence nGPCR-**seq59** <SEQ ID NO. 69> was identified in *H. sapiens*:

TACAGGCCTGAGCATGCTGGGCTCCATCAGCACCAAGCACTGCCTGTCCATCCTGTGGCCCATCTAGTACCGC
TGCCACCACCCCACACACCTGTCAGCAGTCGTGTGTCCTGCTCTGGGCCCTGTCCCTGCTGCAGAGCATCCTG
GAATGGATGTTCTGTGGCTTCCTGTCTAGTGGTGCTGATTCTGTTTGGTGTGAAACATCAGATTTCATCACAG
TCACATGGCTGATTTTTTTATGTGTGGTTCTCTGCGGGTCCAGCCCGGTTCTGCTGGTCAGGATCCTTTGTGG
ATCCCGGAAGATGCCCTTGACCAGGCTGTACATGACCATCCTGCTCAGAGTGCTGGTCTTCCTCCTCTGTGAC
CTGCCCTTTGGCATTCAGTGATTCCTATTTTTCTGGATCCACGTGGATTTGTCACGTTCGTCTAGTTTCCATT
TTCCTGTCCACTCTTAACAGCAGTGCCAACCCCATTATTTACTTCTTCATGGGCTCCTTTAGGCAGCTTCAAA
ACAGGAAGACTCTCTAGCTGGTTCTCCAGAGGGCTCTGCAGGACACGCCTGAGGTGGAAGAAGGCAGATGGCG
GCTTTCTGAGGAAACCCTGGAGCTGTCATGAAGCAGATTGGGGCCATGAGGAAGAGCCTCTGCCCTGTCAGTC
AG

The following amino acid sequence <SEQ ID NO. 70> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 69:

YRPEHAGLHQHQALPVHPVAHLVPLPPPHTPVSSRVSCSGPCPCCRASWNGCSVASCLVVLILFGVKHQISSQ
SHGFFYVWFSAGPARFCWSGSFVDPGRCPPGCTPSCSECWSSSSVTCPLAFSDSYFSGSTWICHVRLVSIFLS
TLNSSANPIIYFFMGSFRQLQNRKTLLVLQRALQDTPEVEEGRWRLSEETLELSSRLGPGRASALSV

The following DNA sequence nGPCR-**seq60** <SEQ ID NO. 71> was identified in *H. sapiens*:

ATGCCGAAGGCAGGCCGCAGAAGAGAAGAGGAGGACGGTGAGGAGGATGAGCCCAGGGAAGCCCCGGGGTGGG
GGCCGCTGGGGGCCTCGCTCCACCCGCAGCAGCAGCATAAGGCTGGCCCCACACATGGTGCAACACAGCAGAG
CCAGCAGCACCGCTGCCACCAGCCACAGCGTCCGGCACAAGTGGCGGCTGGGCTCCCCGAAGAACTGGGTGCA
GGCGCCCGCTGAGCAGCAGGTGCAGCAGCAGGCAGAGGGCCCAGGTGAGGGCGCACACACAGGTGGTCAGGTGG
CGTGGGCGGCGGCACGAGTACCAGGCTGGGAAGAGGGCGGCCAGGCACTGCTCCACGCTGACGGCCGCCAGGA
GACTCAGGCCCACGATGTAGCAGAAGAAGCGCAGCGTTGCCAGGCTGGTCTGCACGAAGCCCGGGAAGTCCAG
CCGGCCTTGCAGCAAGTCGGGGACGATGGCCACCATGTGGCAGCCAAGGAAGATGAGATCCGCGCAGGCCACG
TCCAGGAGGTAGATGGCGAAAGGGTTTCTGTAGACATTGGAGCTGAGC

The following amino acid sequence <SEQ ID NO. 72> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 71:

LSSNVYRNPFAIYLLDVACADLIFLGCHMVAIVPDLLQGRLDFPGFVQTSLATLRFFCYIVGLSLLA
AVSVEQCLAALFPAWYSCRRPRHLTTCVCALTWALCLLLHLTTCVCALTWALCLLLHLLLSGACTLL
LSGACTQFFGEPSRHLCRTLWLVAAVLLALLCCTMCGASLMLLLRVERGPQRPPPRGFPGLILLTVL
LFSSAACLRH

The following DNA sequence **nGPCR-1** <SEQ ID NO. 73> was identified in *H. sapiens*:

ATGGAATCATCTTTCTCATTTGGAGTGATCCTTGCTGTCCTGGCCTCCCTCATCATTGCTACTAACACACTAG
TGGCTGTGGCTGTGCTGCTGTTGATCCACAAGAATGATGGTGTCAGTCTCTGCTTCACCTTGAATCTGGCTGT
GGCTGACACCTTGATTGGTGTGGCCATCTCTGGCCTACTCACAGACCAGCTCTCCAGCCCTTCTCGGCCCACA
CAGAAGACCCTGTGCAGCCTGCGGATGGCATTTGTCACTTCCTCCGCAGCTGCCTCTGTCCTCACGGTCATGC
TGATCACCTTTGACAGGTACCTTGCCATCAAGCAGCCCTTCCGCTACTTGAAGATCATGAGTGGGTTCGTGGC
CGGGGCCTGCATTGCCGGGCTGTGGTTAGTGTCTTACCTCATTGGCTTCCTCCCACTCGGAATCCCCATGTTC
CAGCAGACTGCCTACAAAGGGCAGTGCAGCTTCTTTGCTGTATTTCACCCTCACTTCGTGCTGACCCTCTCCT
GCGTTGGCTTCTTCCCAGCCATGCTCCTCTTTGTCTTCTTCTACTGCGACATGCTCAAGATTGCCTCCATGCA
CAGCCAGCAGATTCGAAAGATGGAACATGCAGGAGCCATGGCTGGAGGTTATCGATCCCCACGGACTCCCAGC
GACTTCAAAGCTCTCCGTACTGTGTCTGTTCTCATTGGGAGCTTTGCTCTATCCTGGACCCCCTTCCTTATCA
CTGGCATTGTGCAGGTGGCCTGCCAGGAGTGTCACCTCTACCTAGTGCTGGAACGGTACCTGTGGCTGCTCGG

CGTGGGCAACTCCCTGCTCAACCCACTCATCTATGCCTATTGGCAGAAGGAGGTGCGACTGCAGCTCTACCAC
ATGGCCCTAGGAGTGAAGAAGGTGCTCACCTCATTCCTCCTCTTTCTCTCGGCCAGGAATTGTGGCCCAGAGA
GGCCCAGGGAAAGTTCCTGTCACATCGTCACTATCTCCAGCTCAGAGTTTGATGGC**TAA**

The following amino acid sequence <SEQ ID NO. 74> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 73:

MESSFSFGVILAVLASLIIATNTLVAVAVLLLIHKNDGVSLCFTLNLAVADTLIGVAISGLLTDQLSSPSRPT
QKTLCSLRMAFVTSSAAASVLTVMLITFDRYLAIKQPFRYLKIMSGFVAGACIAGLWLVSYLIGFLPLGIPMF
QQTAYKGQCSFFAVFHPHFVLTLSCVGFFPAMLLFVFFYCDMLKIASMHSQQIRKMEHAGAMAGGYRSPRTPS
DFKALRTVSVLIGSFALSWTPFLITGIVQVACQECHLYLVLERYLWLLGVGNSLLNPLIYAYWQKEVRLQLY
HMALGVKKVLTSFLLFLSARNCGPERPRESSCHIVTISSSEFDG

The following DNA sequence TL-**GPCR-seq5** <SEQ ID NO. 75> was identified in *H. sapiens*.

AACTGGAAGGGCAGCCGTCTGCCGCCCACGAACACCTTCTCAAGCACTTTGAGTGACCACGGCTTGCAAGCTG
GTGGCTGGCCCCCCGAGTCCCGGGCTCTGAGGCACGGCCGTCGACTTAAGCGTTGCATCCTGTTACCTGGAGA
CCCTCTGAGCTCTCACCTGCTACTTCTGCCGCTGCTTCTGCACAGAGCCCGGGCGAGGACCCCTCCAGG**ATG**C
AGGTCCCGAACAGCACCGGCCCGGACAACGCGACGCTGCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCT
GCCCGTGGTGTACTCGCTGGTGGCGGCGGTCAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGCTGTGCCGG
CGCATGGGGCCCAGATCCCCGTCGGTCATCTTCATGATCAACCTGAGCGTCACGGACCTGATGCTGGCCAGCG
TGTTGCCTTTCCAAATCTACTACCATTGCAACCGCCACCACTGGGTATTCGGGGTGCTGCTTTGCAACGTGGT
GACCGTGGCCTTTTACGCAAACATGTATTCCAGCATCCTCACCATGACCTGTATCAGCGTGGAGCGCTTCCTG
GGGGTCCTGTACCCGCTCAGCTCCAAGCGCTGGCGCCGCCGTCGTTACGCGGTGGCCGCGTGTGCAGGGACCT
GGCTGCTGCTCCTGACCGCCCTGTCCCCGCTGGCGCGCACCGATCTCACCTACCCGGTGCACGCCCTGGGCAT
CATCACCTGCTTCGACGTCCTCAAGTGGACGATGCTCCCCAGCGTGGCCATGTGGGCCGTGTTCCTCTTCACC
ATCTTCATCCTGCTGTTCCTCATCCCGTTCGTGATCACCGTGGCTTGTTACACGGCCACCATCCTCAAGCTGT
TGCGCACGGAGGAGGCGCACGGCCGGGAGCAGCGGAGGCGCGCGGTGGGCCTGGCCGCGGTGGTCTTGCTGGC
CTTTGTCACCTGCTTCGCCCCCAACAACTTCGTGCTCCTGGCGCACATCGTGAGCCGCCTGTTCTACGGCAAG
AGCTACTACCACGTGTACAAGCTCACGCTGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCGTTTGTTTATT
ACTTTGCGTCCCGGGAATTCCAGCTGCGCCTGCGGGAATATTTGGGCTGCCGCCGGGTGCCCAGAGACACCCT
GGACACGCGCCGCGAGAGCCTCTTCTCCGCCAGGACCACGTCCGTGCGCTCCGAGGCCGGTGCGCACCCTGAA
GGGATGGAGGGAGCCACCAGGCCCGGCCTCCAGAGGCAGGAGAGTGTGTTC**TGA**GTCCCGGGGGCGCAGCTTG
GAGAGCCGGGGGCGCAGCTTGGAGGATCCAGGGGCGCATGGAGAGGCCACGGTGCCAGAGGTTCAGGGAGAAC
AGCTGCGTTGCTCCCAGGCACTGCAGAGGCCCGGTGGGGAAGGGTCTCCAGGCTTTATTCCTCCCAGGCACTG
CAGAGGCACCGGTGAGGAAGGGTCTCCAGGCTTCACTCAGGGTAGAGAAACAAGCAAAGCCCAGCAGCGCACA
GGGTGCTTGTTATCCTGCAGAGGGTGCCTCTGCCTCTCTGTGTCAGGGGACAGCTTGTGTCACCACGCCCGGC
TAATTTTTGTATTTTTTTAGTAGAGCTGGGCTGTCACCCCCGAGCTCCTTAGACACTCCTCACACCTGTCCA
TACCCGAGGATGGATATTCAACCAGCCCCACCGCCTACCCGACTCGGTTTCTGGATATCCTCTGTGGGCGAAC
TGCGAGCCCCATTCCCAGCTCTTCTCCCTGCTGACATCGTCCCTTAGCACACCTGTCCATACCCGAGGATGGA
TATTCAACCAGCCCCACCGCCTACCCGACTCGGTTTCTGGATATCCTCTGTGGGCGAACTGCGAGCCCCATTC
CCAGCTCTTCTCCCTGCTGACATCGTCCCTTAGTTGTGGTTCTGGCCTTCTCCATTCTCCTCCAGGGGTTCTG
GTCTCCGTAGCCCGGTGCACGCCGAAATTTCTGTTTATTTCACTCAGGGGCACTGTGGTTGCTGTGGTTGGAA
TTCTTCTTTCAGAGGAGCGCCTGGGGCTCCTGCAAGTCAGCTACTCTCCGTGCCCACTTCCCCTCACACACAC
ACCCCCCTCGTGCCGAATTC

The following amino acid sequence <SEQ ID NO. 76> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 75.

MQVPNSTGPDNATLQMLRNPAIAVALPVVYSLVAAVSIPGNLFSLWVLCRRMGPRSPSVIFMINLSVTDLMLA
SVLPFQIYYHCNRHHWVFGVLLCNVVTVAFYANMYSSILTMTCISVERFLGVLYPLSSKRWRRRRYAVAACAG
TWLLLLTALSPLARTDLTYPVHALGIITCFDVLKWTMLPSVAMWAVFLFTIFILLFLIPFVITVACYTATILK
LLRTEEAHGREQRRRAVGLAAVVLLAFVTCFAPNNFVLLAHIVSRLFYGKSYYHVYKLTLCLSCLNNCLDPFV
YYFASREFQLRLREYLGCRRVPRDTLDTRRESLFSARTTSVRSEAGAHPEGMEGATRPGLQRQESVF

The following DNA sequence **nGPCR-9** <SEQ ID NO. 77> was identified in *H. sapiens*:

ATGGAGTCGGGGCTGCTGCGGCCGGCGCCGGTGAGCGAGGTCATCGTCCTGCATTACAACTACACCGGCAAGC
TCCGCGGTGCGCGCTACCAGCCGGGTGCCGGCCTGCGCGCCGACGCCGTGGTGTGCCTGGCGGTGTGCGCCTT

CATCGTGCTAGAGAATCTAGCCGTGTTGTTGGTGCTCGGACGCCACCCGCGCTTCCACGCTCCCATGTTCCTG
CTCCTGGGCAGCCTCACGTTGTCGGATCTGCTGGCAGGCGCCGCCTACGCCGCCAACATCCTACTGTCGGGGC
CGCTCACGCTGAAACTGTCCCCCGCGCTCTGGTTCGCACGGGAGGGAGGCGTCTTCGTGGCACTCACTGCGTC
CGTGCTGAGCCTCCTGGCCATCGCGCTGGAGCGCAGCCTCACCATGGCGCGCAGGGGGCCCGCGCCCGTCTCC
AGTCGGGGGCGCACGCTGGCGATGGCAGCCGCGGCCTGGGGCGTGTCGCTGCTCCTCGGGCTCCTGCCAGCGC
TGGGCTGGAATTGCCTGGGTCGCCTGGACGCTTGCTCCACTGTCTTGCCGCTCTACGCCAAGGCCTACGTGCT
CTTCTGCGTGCTCGCCTTCGTGGGCATCCTGGCCGCTATCTGTGCACTCTACGCGCGCATCTACTGCCAGGTA
CGCGCCAACGCGCGGCGCCTGCCGGCACGGCCCGGGACTGCGGGGACCACCTCGACCCGGGCGCGTCGCAAGC
CGCGCTCGCTGGCCTTGCTGCGCACGCTCAGCGTGGTGCTCCTGGCCTTTGTGGCATGTTGGGGCCCCCTCTT
CCTGCTGCTGTTGCTCGACGTGGCGTGCCCGGCGCGCACCTGTCCTGTACTCCTGCAGGCCGATCCCTTCCTG
GGACTGGCCATGGCCAACTCACTTCTGAACCCCATCATCTACACGCTCACCAACCGCGACCTGCGCCACGCGC
TCCTGCGCCTGGTCTGCTGCGGACGCCACTCCTGCGGCAGAGACCCGAGTGGCTCCCAGCAGTCGGCGAGCGC
GGCTGAGGCTTCCGGGGGCCTGCGCCGCTGCCTGCCCCCGGGCCTTGATGGGAGCTTCAGCGGCTCGGAGCGC
TCATCGCCCCAGCGCGACGGGCTGGACACCAGCGGCTCCACAGGCAGCCCCGGTGCACCCACAGCCGCCCGGA
CTCTGGTATCAGAACCGGCTGCAGAC**TGA**

The following amino acid sequence <SEQ ID NO. 78> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 77:

MESGLLRPAPVSEVIVLHYNYTGKLRGARYQPGAGLRADAVVCLAVCAFIVLENLAVLLVLGRHPRFHAPMFL
LLGSLTLSDLLAGAAYAANILLSGPLTLKLSPALWFAREGGVFVALTASVLSLLAIALERSLTMARRGPAPVS
SRGRTLAMAAAAWGVSLLLGLLPALGWNCLGRLDACSTVLPLYAKAYVLFCVLAFVGILAAICALYARIYCQV
RANARRLPARPGTAGTTSTRARRKPRSLALLRTLSVVLLAFVACWGPLFLLLLLDVACPARTCPVLLQADPFL
GLAMANSLLNPIIYTLTNRDLRHALLRLVCCGRHSCGRDPSGSQQSASAAEASGGLRRCLPPGLDGSFSGSER
SSPQRDGLDTSGSTGSPGAPTAARTLVSEPAAD

The following DNA sequence **nGPCR-11** <SEQ ID NO. 79> was identified in *H. sapiens*:

ATGTACAACGGGTCGTGCTGCCGCATCGAGGGGGACACCATCTCCCAGGTGATGCCGCCGCTGCTCATTGTGG
CCTTTGTGCTGGGCGCACTAGGCAATGGGGTCGCCCTGTGTGGTTTCTGCTTCCACATGAAGACCTGGAAGCC
CAGCACTGTTTACCTTTTCAATTTGGCCGTGGCTGATTTCCTCCTTATGATCTGCCTGCCTTTTCGGACAGAC
TATTACCTCAGACGTAGACACTGGGCTTTTGGGGACATTCCCTGCCGAGTGGGGCTCTTCACGTTGGCCATGA
ACAGGGCCGGGAGCATCGTGTTCCTTACGGTGGTGGCTGCGGACAGGTATTTCAAAGTGGTCCACCCCCACCA
CGCGGTGAACACTATCTCCACCCGGGTGGCGGCTGGCATCGTCTGCACCCTGTGGGCCCTGGTCATCCTGGGA
ACAGTGTATCTTTTGCTGGAGAACCATCTCTGCGTGCAAGAGACGGCCGTCTCCTGTGAGAGCTTCATCATGG
AGTCGGCCAATGGCTGGCATGACATCATGTTCCAGCTGGAGTTCTTTATGCCCCTCGGCATCATCTTATTTTG
CTCCTTCAAGATTGTTTGGAGCCTGAGGCGGAGGCAGCAGCTGGCCAGACAGGCTCGGATGAAGAAGGCGACC
CGGTTCATCATGGTGGTGGCAATTGTGTTCATCACATGCTACCTGCCCAGCGTGTCTGCTAGACTCTATTTCC
TCTGGACGGTGCCCTCGAGTGCCTGCGATCCCTCTGTCCATGGGGCCCTGCACATAACCCTCAGCTTCACCTA
CATGAACAGCATGCTGGATCCCCTGGTGTATTATTTTTCAAGCCCCTCCTTTCCCAAATTCTACAACAAGCTC
AAAATCTGCAGTCTGAAACCCAAGCAGCCAGGACACTCAAAAACACAAAGGCCGGAAGAGATGCCAATTTCGA
ACCTCGGTCGCAGGAGTTGCATCAGTGTGGCAAATAGTTTCCAAAGCCAGTCTGATGGGCAATGGGATCCCCA
CATTGTTGAGTGGCAC**TGA**

The following amino acid sequence <SEQ ID NO. 80> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 79:

MYNGSCCRIEGDTISQVMPPLLIVAFVLGALGNGVALCGFCFHMKTWKPSTVYLFNLAVADFLLMICLPFRTD
YYLRRRHWAFGDIPCRVGLFTLAMNRAGSIVFLTVVAADRYFKVVHPHHAVNTISTRVAAGIVCTLWALVILG
TVYLLLENHLCVQETAVSCESFIMESANGWHDIMFQLEFFMPLGIILFCSFKIVWSLRRRQQLARQARMKKAT
RFIMVVAIVFITCYLPSVSARLYFLWTVPSSACDPSVHGALHITLSFTYMNSMLDPLVYYFSSPSFPKFYNKL
KICSLKPKQPGHSKTQRPEEMPISNLGRRSCISVANSFQSQSDGQWDPHIVEWH

The following DNA sequence **nGPCR-16** <SEQ ID NO. 81> was identified in *H. sapiens*:

ATGACAGGTGACTTCCCAAGTATGCCTGGCCACAATACCTCCAGGAATTCCTCTTGCGATCCTATAGACACCC
CACTTAATCAGCCTCTACTTCATAGTGCTTATTGGCGGGCTGGTGGGTGTCATTTCCATTCTTTTCCTCCTGG
TGAAAATGAACACCCGGTCAGTGACCACCATGGCGGTCATTAACTTGGTGGTGGTCCACAGCGTTTTTCTGCT
GACAGTGCCATTTCGCTTGACCTACCTCATCAAGAAGACTTGGATGTTTGGGCTGCCCTTCTGCAAATTTGTG

AGTGCCATGCTGCACATCCACATGTACCTCACGTTCCTATTCTATGTGGTGATCCTGGTCACCAGATACCTCA
TCTTCTTCAAGTGCAAAGACAAAGTGGAATTCTACAGAAAACTGCATGCTGTGGCTGCCAGTGCTGGCATGTG
GACGCTGGTGATTGTCATTGTGGTACCCCTGGTTGTCTCCCGGTATGGAATCCATGAGGAATACAATGAGGAG
CACTGTTTTAAATTTCACAAAGAGCTTGCTTACACATATGTGAAAATCATCAACTATATGATAGTCATTTTTG
TCATAGCCGTTGCTGTGATTCTGTTGGTCTTCCAGGTCTTCATCATTATGTTGATGGTGCAGAAGCTACGCCA
CTCTTTACTATCCCACCAGGAGTTCTGGGCTCAGCTGAAAAACCTATTTTTTATAGGGGTCATCCTTGTTTGT
TTCCTTCCCTACCAGTTCTTTAGGATCTATTACTTGAATGTTGTGACGCATTCCAATGCCTGTAACAGCAAGG
TTGCATTTTATAACGAAATCTTCTTGAGTGTAACAGCAATTAGCTGCTATGATTTGCTTCTCTTTGTCTTTGG
GGGAAGCCATTGGTTTAAGCAAAAGATAATTGGCTTATGGAATTGTGTTTTGTGCCGT**TAG**CCACAAACTACA
GTATTCATATTTGCTTCCTTTATATTGGGAATAAAAATGGGTATAGGGGAGGTAAGAATGGTATTTCATTACT
TGATCAAAACCATGCCTTGATGTACCCAAAACAAAAGGACTATAAAATGCAAGAGCCCTCATTGTAGTCCTTA
TGGGATCCCTCCCATCTCTGAGTGATGGCCGTACAAAGACCAGTGTTGTTGAATCCACCTGGAGTTGCAATAT
TACATTATTTTCCAGTACAGAATGTCTGTGTGGCCCATGAAAGCAACATAGGTTTTAAGAGTTTTAGAGTTTC
ATTAGCTCATTCTAAGTTCCTCTGTTTGAAGCATGGTCTCTTAGGTTTTGGACTGAACTCAGACCTTTAGTTC
TTTTCATCCCACTTCACCTTAGGTAAGTAAATTCTGGCCACCACCCAGCTCCAAAGACACAAACTCTCCTTCG
CTAACCAGGTTAGATGTCCCATTCATCTCATGCCCTGATAAAAACTGATAAGGGGAGAGAATAGTTAAAAATT
TTTCTAGGGTATCATAACTCTGGTAGGAAGTCATCTGTCTAGAAATCAAGAGAAAAAGAACGTGTGGCCTCCT
GTTATAACAAGGGTTTCTAGATTTGTCCTGTGAAAGGTCGTTTAAGGACTTGGGGATCAACTTCCTCAATTAT
CACCAATTGCACTGTTGCTCCAAAAATCATTTAAAAGCTTACTGGACATATCTACATAATGGTGAAACTGTAA
TTTAGAGACTATCCCTGACTAATGTGCTGGTAGGCATTAAAATGAGTTCCCAAGGGAAGTGATTAAAATTTTT
TTCTCTTCTGTTTTTTGAGAGAATTTCTAGATGTCCTGGGCCACAGTTAATTAAGATTTTTAGGGGGGACAGA
AAGTTATACTGAAATCTTTAGAGCTCCCTTCCGCCGTTAAAATTATATATATATATATTTAAATTATACCTTA
AGTTCTGGGGTACATGTGCAGAATGTGCAGGTTTGTTACATAGGTATACACGTGCCATGGTGGTTTGCGGCAC
CTGTCAACCCATCTACATTAGGTATTTCTCCTAATGCTCTCCCTCCCCTAGCCCCCCACCCCTGGACAGGCCC
CATTGTGTGATGTTCCCCTCCCTGTGTCCATGTGTTTTCATTGTTCAACTCCCACTTCTAAGTGAGAACATGC
GGTGTTTGGTTTTCTGTTCCTGTGTTAGTTTGCTGAGAATGATGGTTTCCAGGTTAAAATTATATATTTTTAA
ATAAATGAAAACTGTGTTTTTAAAAGAGGACTTTTGAGAAGTATATAGAAAAACCATTAATTTAGACTCTGTG
AGATTAGGTTGCATGAAGAAGGTTTTCTGAATATTTGAAGAGTGGATAAATAAATGTCCCCCAAAGCAATAAA
ATCATAATCCTTTAAAATATAGGAAAAATAACTAATGGGAACTAGGCTTAATACTCGGGATGAAATAATCTGT
ACAACAAACTCCCATGACACATGTTTACCTATGTAACAAACCTGCACATGTACCCCTGAACTTAAAATAAAAT
TTAAAGTATAATAATAAAATAATATGGATTTTCTTT

The following amino acid sequence <SEQ ID NO. 82> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 81:

MTGDFPSMPGHNTSRNSSCDPIVTPHLISLYFIVLIGGLVGVISILFLLVKMNTRSVTTMAVINLVVVHSVFL
LTVPFRLTYLIKKTWMFGLPFCKFVSAMLHIHMYLTFLFYVVILVTRYLIFFKCKDKVEFYRKLHAVAASAGM
WTLVIVIVVPLVVSRYGIHEEYNEEHCFKFHKELAYTYVKIINYMIVIFVIAVAVILLVFQVFIIMLMVQKLR
HSLLSHQEFWAQLKNLFFIGVILVCFLPYQFFRIYYLNVVTHSNACNSKVAFYNEIFLSVTAISCYDLLLFVF
GGSHWFKQKIIGLWNCVLCR

The following DNA sequence **nGPCR-40** <SEQ ID NO. 83> was identified in *H. sapiens*:

GCAGGAGCACTGAAAATCAGGAACAATCCTGTATTTTTTGTGATAATCAACAAGGACAAAACTTCTCCATATG
TAAATAACAGCGTT**ATG**AGCAGCAATTCATCCCTGCTGGTGGCTGTGCAGCTGTGCTACGCGAACGTGAATGG
GTCCTGTGTGAAAATCCCCTTCTCGCCGGGATCCCGGGTGATTCTGTACATAGTGTTTGGCTTTGGGGCTGTG
CTGGCTGTGTTTGGAAACCTCCTGGTGATGATTTCAATCCTCCATTTCAAGCAGCTGCACTCTCCGACCAATT
TTCTCGTTGCCTCTCTGGCCTGCGCTGATTTCTTGGTGGGTGTGACTGTGATGCCCTTCAGCATGGTCAGGAC
GGTGGAGAGCTGCTGGTATTTTGGGAGGAGTTTTTGTACTTTCCACACCTGCTGTGATGTGGCATTTTGTTAC
TCTTCTCTCTTTCACTTGTGCTTCATCTCCATCGACAGGTACATTGCGGTTACTGACCCCCTGGTCTATCCTA
CCAAGTTCACCGTATCTGTGTCAGGAATTTGCATCAGCGTGTCCTGGATCCTGCCCCTCATGTACAGCGGTGC
TGTGTTCTACACAGGTGTCTATGACGATGGGCTGGAGGAATTATCTGATGCCCTAAACTGTATAGGAGGTTGT
CAGACCGTTGTAAATCAAAACTGGGTGTTGACAGATTTTCTATCCTTCTTTATACCTACCTTTATTATGATAA
TTCTGTATGGTAACATATTTCTTGTGGCTAGACGACAGGCGAAAAAGATAGAAAATACTGGTAGCAAGACAGA
ATCATCCTCAGAGAGTTACAAAGCCAGAGTGGCCAGGAGAGAGAGAAAAGCAGCTAAAACCCTGGGGGTCACA
GTGGTAGCATTTATGATTTCATGGTTACCATATAGCATTGATTCATTAATTGATGCCTTTATGGGCTTTATAA
CCCCTGCCTGTATTTATGAGATTTGCTGTTGGTGTGCTTATTATAACTCAGCCATGAATCCTTTGATTTATGC
TTTATTTTACCCATGGTTTAGGAAAGCAATAAAAGTTATTGTAACTGGTCAGGTTTTAAAGAACAGTTCAGCA
ACCATGAATTTGTTTTCTGAACATATA**TAA**

The following amino acid sequence <SEQ ID NO. 84> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 83:

MSSNSSLLVAVQLCYANVNGSCVKIPFSPGSRVILYIVFGFGAVLAVFGNLLVMISILHFKQLHSPTNFLVAS
LACADFLVGVTVMPFSMVRTVESCWYFGRSFCTFHTCCDVAFCYSSLFHLCFISIDRYIAVTDPLVYPTKFTV
SVSGICISVSWILPLMYSGAVFYTGVYDDGLEELSDALNCIGGCQTVVNQNWVLTDFLSFFIPTFIMIILYGN
IFLVARRQAKKIENTGSKTESSSESYKARVARRERKAAKTLGVTVVAFMISWLPYSIDSLIDAFMGFITPACI
YEICCWCAYYNSAMNPLIYALFYPWFRKAIKVIVTGQVLKNSSATMNLFSEHI

The following DNA sequence **nGPCR-54** <SEQ ID NO. 85> was identified in *H. sapiens*:

ACC**ATG**AATGAGCCACTAGACTATTTAGCAAATGCTTCTGATTTCCCCGATTATGCAGCTGCTTTTGGAAATT
GCACTGATGAAAACATCCCACTCAAGATGCACTACCTCCCTGTTATTTATGGCATTATCTTCCTCGTGGGATT
TCCAGGCAATGCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGGAAGAGCAGCACCATCATTATG
CTGAACCTGGCCTGCACAGATCTGCTGTATCTGACCAGCCTCCCCTTCCTGATTCACTACTATGCCAGTGGCG
AAAACTGGATCTTTGGAGATTTCATGTGTAAGTTTATCCGCTTCAGCTTCCATTTCAACCTGTATAGCAGCAT
CCTCTTCCTCACCTGTTTCAGCATCTTCCGCTACTGTGTGATCATTCACCCAATGAGCTGCTTTTCCATTCAC
AAAACTCGATGTGCAGTTGTAGCCTGTGCTGTGGTGTGGATCATTTCACTGGTAGCTGTCATTCCGATGACCT
TCTTGATCACATCAACCAACAGGACCAACAGATCAGCCTGTCTCGACCTCACCAGTTCGGATGAACTCAATAC
TATTAAGTGGTACAACCTGATTTTGACTGCAAGTACTTTCTGCCTCCCCTTGGTGATAGTGACACTTTGCTAT
ACCACGATTATCCACACTTTGACCCATGGACTGCAAACTGACAGCTGCCTTAAGCAGAAAGCACGAAGGCTAA
CCATTCTGCTACTCCTTGCATTTTACGTATGTTTTTACCCTTCCATATCTTGAGGGTCATTCAGGATCGAAT
CTCAGCCTGCTTTCAATCAGTTGTTCCATTGAGAATCAGATCCATGAAGCTTACATCGTTTCTAGACCATTAT
GCTGCTCTGAACACCTTTGGTAACCTGTTACTATATGTGGTGGTCAGCGACAACTTTCAGCAGGCTGTCTGCT
CAACAGTGAGATGCAAAGTAAGCGGGAACCTTGAGCAAGCAAAGAAAATTAGTTACTCAAACAACCCT**TGA**

The following amino acid sequence <SEQ ID NO. 86> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 85:

MNEPLDYLANASDFPDYAAAFGNCTDENIPLKMHYLPVIYGIIFLVGFPGNAVVISTYIFKMRPWKSSTIIML
NLACTDLLYLTSLPFLIHYYASGENWIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRYCVIIHPMSCFSIHK
TRCAVVACAVVWIISLVAVIPMTFLITSTNRTNRSACLDLTSSDELNTIKWYNLILTASTFCLPLVIVTLCYT
TIIHTLTHGLQTDSCLKQKARRLTILLLLAFYVCFLPFHILRVIQDRISACFQSVVPLRIRSMKLTSFLDHYA
ALNTFGNLLLYVVVSDNFQQAVCSTVRCKVSGNLEQAKKISYSN

The following DNA sequence **nGPCR-56** <SEQ ID NO. 87> was identified in *H. sapiens*:

AAAAATTGCTGTACTGAACTATTGAATGGAACTTGGAAATAAAGTCCCTTCCAAAATAACTATTCTTCAACAG
AGAGTAATAGGTAAATGTTTTAGAAGTGAGAGGACTCAAATTGCCAATGATTTACTCTTTTATTTTTCCTCCT
AGGTTTCTGGGATAAGTATGTGCAAATAAAAAATAAACATGAGAAGGAACTGTAACCTGATTATGGATTTGGG
AAAAAGATAAATCAACACACAAAGGGAAAAGTAAACTGATTGACAGCCCTCAGGA**ATG**ATGCCCTTTTGCCAC
AATATAATTAATATTTCCTGTGTGAAAAACAACTGGTCAAATGATGTCCGTGCTTCCCTGTACAGTTTAATGG
TGCTCATAATTCTGACCACACTCGTTGGCAATCTGATAGTTATTGTTTCTATATCACACTTCAAACAACTTCA
TACCCCAACAAATTGGCTCATTCATTCCATGGCCACTGTGGACTTTCTTCTGGGGTGTCTGGTCATGCCTTAC
AGTATGGTGAGATCTGCTGAGCACTGTTGGTATTTTGGAGAAGTCTTCTGTAAAATTCACACAAGCACCGACA
TTATGCTGAGCTCAGCCTCCATTTTCCATTTGTCTTTCATCTCCATTGACCGCTACTATGCTGTGTGTGATCC
ACTGAGATATAAAGCCAAGATGAATATCTTGGTTATTTGTGTGATGATCTTCATTAGTTGGAGTGTCCCTGCT
GTTTTTGCATTTGGAATGATCTTTCTGGAGCTAAACTTCAAAGGCGCTGAAGAGATATATTACAAACATGTTC
ACTGCAGAGGAGGTTGCTCTGTCTTCTTTAGCAAAATATCTGGGGTACTGACCTTTATGACTTCTTTTTATAT
ACCTGGATCTATTATGTTATGTGTCTATTACAGAATATATCTTATCGCTAAAGAACAGGCAAGATTAATTAGT
GATGCCAATCAGAAGCTCCAAATTGGATTGGAAATGAAAAATGGAATTTCACAAAGCAAAGAAAGGAAAGCTG
TGAAGACATTGGGGATTGTGATGGGAGTTTTCCTAATATGCTGGTGCCCTTTCTTTATCTGTACAGTCATGGA
CCCTTTTCTTCACTACATTATTCCACCTACTTTGAATGATGTA

The following amino acid sequence <SEQ ID NO. 88> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 87:

MMPFCHNIINISCVKNNWSNDVRASLYSLMVLIILTTLVGNLIVIVSISHFKQLHTPTNWLIHSMATVDFLLG
CLVMPYSMVRSAEHCWYFGEVFCKIHTSTDIMLSSASIFHLSFISIDRYYAVCDPLRYKAKMNILVICVMIFI

SWSVPAVFAFGMIFLELNFKGAEEIYYKHVHCRGGCSVFFSKISGVLTFMTSFYIPGSIMLCVYYRIYLIAKE
QARLISDANQKLQIGLEMKNGISQSKERKAVKTLGIVMGVFLICWCPFFICTVMDPFLHYIIPPTLNDARGSR
ANSA

The following DNA sequence **nGPCR-56** <SEQ ID NO. 89> was identified in *H. sapiens*:

GGA**ATG**ATGCCCTTTTGCCACAATATAATTAATATTTCCTGTGTGAAAAACAACTGGTCAAATGATGTCCGTG
CTTCCCTGTACAGTTTAATGGTGCTCATAATTCTGACCACACTCGTTGGCAATCTGATAGTTATTGTTTCTAT
ATCACACTTCAAACAACTTCATACCCCAACAAATTGGCTCATTCATTCCATGGCCACTGTGGACTTTCTTCTG
GGGTGTCTGGTCATGCCTTACAGTATGGTGAGATCTGCTGAGCACTGTTGGTATTTTGGAGAAGTCTTCTGTA
AAATTCACACAAGCACCGACATTATGCTGAGCTCAGCCTCCATTTTCCATTTGTCTTTCATCTCCATTGACCG
CTACTATGCTGTGTGTGATCCACTGAGATATAAAGCCAAGATGAATATCTTGGTTATTTGTGTGATGATCTTC
ATTAGTTGGAGTGTCCCTGCTGTTTTTGCATTTGGAATGATCTTTCTGGAGCTAAACTTCAAAGGCGCTGAAG
AGATATATTACAAACATGTTCACTGCAGAGGAGGTTGCTCTGTCTTCTTTAGCAAAATATCTGGGGTACTGAC
CTTTATGACTTCTTTTTATATACCTGGATCTATTATGTTATGTGTCTATTACAGAATATATCTTATCGCTAAA
GAACAGGCAAGATTAATTAGTGATGCCAATCAGAAGCTCCAAATTGGATTGGAAATGAAAAATGGAATTTCAC
AAAGCAAAGAAAGGAAAGCTGTGAAGACATTGGGGATTGTGATGGGAGTTTTCCTAATATGCTGGTGCCCTTT
CTTTATCTGTACAGTCATGGACCCTTTTCTTCACTACATTATTCCACCTACTTTGAATGATGTATTGATTTGG
TTTGGCTACTTGAACTCTACATTTAATCCAATGGTTTATGCATTTTTCTATCCTTGGTTTAGAAAAGCACTGA
AGATGATGCTGTTTGGTAAAATTTTCCAAAAAGATTCATCCAGGTGTAAATTATTTTTGGAATTGAGTTCA**TA
G**

The following amino acid sequence <SEQ ID NO. 90> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 89:

MMPFCHNIINISCVKNNWSNDVRASLYSLMVLIILTTLVGNLIVIVSISHFKQLHTPTNWLIHSMATVDFLLG
CLVMPYSMVRSAEHCWYFGEVFCKIHTSTDIMLSSASIFHLSFISIDRYYAVCDPLRYKAKMNILVICVMIFI
SWSVPAVFAFGMIFLELNFKGAEEIYYKHVHCRGGCSVFFSKISGVLTFMTSFYIPGSIMLCVYYRIYLIAKE
QARLISDANQKLQIGLEMKNGISQSKERKAVKTLGIVMGVFLICWCPFFICTVMDPFLHYIIPPTLNDVLIWF
GYLNSTFNPMVYAFFYPWFRKALKMMLFGKIFQKDSSRCKLFLELSS

The following DNA sequence **nGPCR-58** <SEQ ID NO. 91> was identified in *H. sapiens*:

CTGTAAAGTAGATTGTATGAGGACTCCATGAGGTCATCCACTTCAAGTCCTTGGCATAGGATAATTACTCAAA
AGGTGATGACAATGGCGCAGGGAGGGATGGTGACTTGCCTGGAGATGCACAGCACCGTCTCTCCCATACTCGG
TCATTCACACCATCATTGATTCACCAGGCACCACTCCGTGTCCAGCAGGACTCTGGGGACCCCAAATGGACAC
TACC**ATG**GAAGCTGACCTGGGTGCCACTGGCCACAGGCCCCGCACAGAGCTTGATGATGAGGACTCCTACCCC
CAAGGTGGCTGGGACACGGTCTTCCTGGTGGCCCTGCTGCTCCTTGGGCTGCCAGCCAATGGGTTGATGGCGT
GGCTGGCCGGCTCCCAGGCCCGGCATGGAGCTGGCACGCGTCTGGCGCTGCTCCTGCTCAGCCTGGCCCTCTC
TGACTTCTTGTTCCTGGCAGCAGCGGCCTTCCAGATCCTAGAGATCCGGCATGGGGGACACTGGCCGCTGGGG
ACAGCTGCCTGCCGCTTCTACTACTTCCTATGGGGCGTGTCCTACTCCTCCGGCCTCTTCCTGCTGGCCGCCC
TCAGCCTCGACCGCTGCCTGCTGGCGCTGTGCCCACACTGGTACCCTGGGCACCGCCCAGTCCGCCTGCCCCT
CTGGGTCTGCGCCGGTGTCTGGGTGCTGGCCACACTCTTCAGCGTGCCCTGGCTGGTCTTCCCCGAGGCTGCC
GTCTGGTGGTACGACCTGGTCATCTGCCTGGACTTCTGGGACAGCGAGGAGCTGTCGCTGAGGATGCTGGAGG
TCCTGGGGGGCTTCCTGCCTTTCCTCCTGCTGCTCGTCTGCCACGTGCTCACCCAGGCCACAGCCTGTCGCAC
CTGCCACCGCCAACAGCAGCCCGCAGCCTGCCGGGGCTTCGCCCGTGTGGCCAGGACCATTCTGTCAGCCTAT
GTGGTCCTGAGGCTGCCCTACCAGCTGGCCCAGCTGCTCTACCTGGCCTTCCTGTGGGACGTCTACTCTGGCT
ACCTGCTCTGGGAGGCCCTGGTCTACTCCGACTACCTGATCCTACTCAACAGCTGCCTCAGCCCCTTCCTCTG
CCTCATGGCCAGTGCCGACCTCCGGACCCTGCTGCGCTCCGTGCTCTCGTCCTTCGCGGCAGCTCTCTGCGAG
GAGCGGCCGGGCAGCTTCACGCCCACTGAGCCACAGACCCAGCTAGATTCTGAGGGTCCAACTCTGCCAGAGC
CGATGGCAGAGGCCCAGTCACAGATGGATCCTGTGGCCCAGCCTCAGGTGAACCCCACACTCCAGCCACGATC
GGATCCCACAGCTCAGCCACAGCTGAACCCTACGGCCCAGCCACAGTCGGATCCCACAGCCCAGCCACAGCTG
AACCTCATGGCCCAGCCACAGTCAGATTCTGTGGCCCAGCCACAGGCAGACACTAACGTCCAGACCCCTGCAC
CTGCTGCCAGTTCTGTGCCCAGTCCCTGTGATGAAGCTTCCCCAACCCCATCCTCGCATCCTACCCCAGGGGC
CCTTGAGGACCCAGCCACACCTCCTGCCTCTGAAGGAGAAAGCCCCAGCAGCACCCCGCCAGAGGCGGCCCCG
GGCGCAGGCCCCACG**TGA**GGGTCCAGGAACACGCAGGCCCACCAGAGCAGTGAAAGAGCCCAGGGCAGACAGA
GGAACCAGCCAGTCAGA

The following amino acid sequence <SEQ ID NO. 92> is the predicted amino

acid sequence derived from the DNA sequence of SEQ ID NO. 91:

LAWRCTAPSLPYSVIHTIIDSPGTTPCPAGLWGPQMDTTMEADLGATGHRPRTELDDEDSYPQGGWDTVFLVA
LLLLGLPANGLMAWLAGSQARHGAGTRLALLLLSLALSDFLFLAAAAFQILEIRHGGHWPLGTAACRFYYFLW
GVSYSSGLFLLAALSLDRCLLALCPHWYPGHRPVRLPLWVCAGVWVLATLFSVPWLVFPEAAVWWYDLVICLD
FWDSEELSLRMLEVLGGFLPFLLLLVCHVLTQATACRTCHRQQQPAACRGFARVARTILSAYVVLRLPYQLAQ
LLYLAFLWDVYSGYLLWEALVYSDYLILLNSCLSPFLCLMASADLRTLLRSVLSSFAAALCEERPGSFTPTEP
QTQLDSEGPTLPEPMAEAQSQMDPVAQPQVNPTLQPRSDPTAQPQLNPTAQPQSDPTAQPQLNLMAQPQSDSV
AQPQADTNVQTPAPAASSVPSPCDEASPTPSSHPTPGALEDPATPPASEGESPSSTPPEAAPGAGPT

The following DNA sequence **nGPCR-58** <SEQ ID NO. 93> was identified in *H. sapiens*:

ATGGACACTACCATGGAAGCTGACCTGGGTGCCACTGGCCACAGGCCCCGCACAGAGCTTGATGATGAGGACT
CCTACCCCCAAGGTGGCTGGGACACGGTCTTCCTGGTGGCCCTGCTGCTCCTTGGGCTGCCAGCCA
ATGGGTTGATGGCGTGGCTGGCCGGCTCCCAGGCCCGGCATGGAGCTGGCACGCGTCTGGCGCTGCTCCTGCT
CAGCCTGGCCCTCTCTGACTTCTTGTTCCTGGCAGCAGCGGCCTTCCAGATCCTAGAGATCCGGCATGGGGGA
CACTGGCCGCTGGGGACAGCTGCCTGCCGCTTCTACTACTTCCTATGGGGCGTGTCCTACTCCTCCGGCCTCT
TCCTGCTGGCCGCCCTCAGCCTCGACCGCTGCCTGCTGGCGCTGTGCCCACACTGGTACCCTGGGCACCGCCC
AGTCCGCCTGCCCCTCTGGGTCTGCGCCGGTGTCTGGGTGCTGGCCACACTCTTCAGCGTGCCCTGGCTGGTC
TTCCCCGAGGCTGCCGTCTGGTGGTACGACCTGGTCATCTGCCTGGACTTCTGGGACAGCGAGGAGCTGTCGC
TGAGGATGCTGGAGGTCCTGGGGGGCTTCCTGCCTTTCCTCCTGCTGCTCGTCTGCCACGTGCTCACCCAGGC
CACAGCCTGTCGCACCTGCCACCGCCAACAGCAGCCCGCAGCCTGCCGGGGCTTCGCCCGTGTGGCCAGGACC
ATTCTGTCAGCCTATGTGGTCCTGAGGCTGCCCTACCAGCTGGCCCAGCTGCTCTACCTGGCCTTCCTGTGGG
ACGTCTACTCTGGCTACCTGCTCTGGGAGGCCCTGGTCTACTCCGACTACCTGATCCTACTCAACAGCTGCCT
CAGCCCCTTCCTCTGCCTCATGGCCAGTGCCGACCTCCGGACCCTGCTGCGCTCCGTGCTCTCGTCCTTCGCG
GCAGCTCTCTGCGAGGAGCGGCCGGGCAGCTTCACGCCCACTGAGCCACAGACCCAGCTAGATTCTGAGGGTC
CAACTCTGCCAGAGCCGATGGCAGAGGCCCAGTCACAGATGGATCCTGTGGCCCAGCCTCAGGTGAACCCCAC
ACTCCAGCCACGATCGGATCCCACAGCTCAGCCACAGCTGAACCCTACGGCCCAGCCACAGTCGGATCCCACA
GCCCAGCCACAGCTGAACCTCATGGCCCAGCCACAGTCAGACTCTGTGGCCCAGCCACAGGCAGACACTAACG
TCCAGACCCCTGCACCTGCTGCCAGTTCTGTGCCCAGTCCCTGTGATGAAGCTTCCCCAACCCCATCCTCGCA
TCCTACCCCAGGGGCCCTTGAGGACCCAGCCACACCTCCTGCCTCTGAAGGAGAAAGCCCCAGCAGCACCCCG
CCAGAGGCGGCCCCGGGCGCAGGCCCCACG**TGA**

The following amino acid sequence <SEQ ID NO. 94> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 93:

MDTTMEADLGATGHRPRTELDDEDSYPQGGWDTVFLVALLLLGLPANGLMAWLAGSQARHGAGTRLALLLLSL
ALSDFLFLAAAAFQILEIRHGGHWPLGTAACRFYYFLWGVSYSSGLFLLAALSLDRCLLALCPHWYPGHRPVR
LPLWVCAGVWVLATLFSVPWLVFPEAAVWWYDLVICLDFWDSEELSLRMLEVLGGFLPFLLLLVCHVLTQATA
CRTCHRQQQPAACRGFARVARTILSAYVVLRLPYQLAQLLYLAFLWDVYSGYLLWEALVYSDYLILLNSCLSP
FLCLMASADLRTLLRSVLSSFAAALCEERPGSFTPTEPQTQLDSEGPTLPEPMAEAQSQMDPVAQPQVNPTLQ
PRSDPTAQPQLNPTAQPQSDPTAQPQLNLMAQPQSDSVAQPQADTNVQTPAPAA

The following DNA sequence nGPCR-3 <SEQ ID NO. 185> was identified in *H. sapiens*:

AGGCTCGCGCCCGAAGCAGAGCCATGAGAACCCCAGGGTGCCTGGCGAGCCGCTAGCGCC**ATG**GGCCCCGGCG
AGGCGCTGCTGGCGGGTCTCCTGGTGATGGTACTGGCCGTGGCGCTGCTATCCAACGCACTGGTGCTGCTTTG
TTGCGCCTACAGCGCTGAGCTCCGCACTCGAGCCTCAGGCGTCCTCCTGGTGAATCTGTCTCTGGCCACCTG
CTGCTGGCGGCGCTGGACATGCCCTTCACGCTGCTCGGTGTGATGCGCGGGCGGACACCGTCGGCGCCCGGCG
CATGCCAAGTCATTGGCTTCCTGGACACCTTCCTGGCGTCCAACGCGGCGCTGAGCGTGGCGGCGCTGAGCGC
AGACCAGTGGCTGGCAGTGGGCTTCCCACTGCGCTACGCCGGACGCCTGCGACCGCGCTATGCCGGCCTGCTG
CTGGGCTGTGCCTGGGGACAGTCGCTGGCCTTCTCAGGCGCTGCACTTGGCTGCTCGTGGCTTGGCTACAGCA
GCGCCTTCGCGTCCTGTTCGCTGCGCCTGCCGCCCGAGCCTGAGCGTCCGCGCTTCGCAGCCTTCACCGCCAC
GCTCCATGCCGTGGGCTTCGTGCTGCCGCTGGCGGTGCTCTGCCTCACCTCGCTCCAGGTGCACCGGGTGGCA
CGCAGACACTGCCAGCGCATGGACACCGTCACCATGAAGGCGCTCGCGCTGCTCGCCGACCTGCACCCCAGTG
TGCGGCAGCGCTGCCTCATCCAGCAGAAGCGGCGCCGCCACCGCGCCACCAGGAAGATTGGCATTGCTATTGC
GACCTTCCTCATCTGCTTTGCCCCGTATGTCATGACCAGGCTGGCGGAGCTCGTGCCCTTCGTCACCGTGAAC
GCCCAGTGGGGCATCCTCAGCAAGTGCCTGACCTACAGCAAGGCGGTGGCCGACCCGTTCACGTACTCTCTGC
TCCGCCGGCCGTTCCGCCAAGTCCTGGCCGGCATGGTGCACCGGCTGCTGAAGAGAACCCCGCGCCCAGCATC
CACCCATGACAGCTCTCTGGATGTGGCCGGCATGGTGCACCAGCTGCTGAAGAGAACCCCGCGCCCAGCGTCC

ACCCACAACGGCTCTGTGGACACAGAGAATGATTCCTGCCTGCAGCAGACACAC**TGA**GGGCCTGGCAGGGCTC
ATCGCCCCCACCTTCTAAGA

The following amino acid sequence <SEQ ID NO. 186> is the predicted amino acid sequence derived from the DNA sequence of SEQ ID NO. 185:

MGPGEALL**AGLLVMVLAVALLSNALVLLCCA**YSAELRTRASG**VLLVNLSLGHLLLAALDMPFTLL**GVMRGRTP
SAPGAC**QVIGFLDTFLASNAALSVAALSA**DQWLAVGFPLRYAGRLRPR**YAGLLLGCAWGQSLAFSGAALGC**SW
LGYSSAFASCSLRLPPEPERPR**FAAFTATLHAVGFVLPLAVLCLT**SLQVHRVARRHCQRMDTVTMKALALLAD
LHPSVRQRCLIQQKRRRHRATR**KIGIAIATFLICFAPYVMT**RLAELVPFVTVNAQWGILSK**CLTYSKAVADPF**
TYSLLRRPFRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVHQLLKRTPRPASTHNGSVDTENDSCLQQTH

EXAMPLE 2: CLONING OF nGPCR-X

To isolate a cDNA clone encoding full length nGPCR-x, a DNA fragment corresponding to a nucleotide sequence set forth in odd numbered nucleotide sequences ranging from SEQ ID NO: 1-93, or a portion thereof, can be used as a probe for hybridization screening of a phage cDNA library. The DNA fragment is amplified by the polymerase chain reaction (PCR) method. The PCR reaction mixture of 50 μl contains polymerase mixture (0.2 mM dNTPs, 1x PCR Buffer and 0.75 μl Expand High Fidelity Polymerase (Roche Biochemicals)), 1 μg of plasmid, and 50 pmoles of forward primer and 50 pmoles of reverse primer. The primers are preferably 10 to 25 nucleotides in length and are determined by procedures well known to those skilled in the art. Amplification is performed in an Applied Biosystems PE2400 thermocycler, using the following program: 95°C for 15 seconds, 52°C for 30 seconds and 72°C for 90 seconds; repeated for 25 cycles. The amplified product is separated from the plasmid by agarose gel electrophoresis, and purified by Qiaquick™ gel extraction kit (Qiagen).

A lambda phage library containing cDNAs cloned into lambda ZAPII phage-vector is plated with *E. coli* XL-1 blue host, on 15 cm LB-agar plates at a density of 50,000 pfu per plate, and grown overnight at 37°C; (plated as described by Sambrook *et al.*, supra). Phage plaques are transferred to nylon membranes (Amersham Hybond NJ), denatured for 2 minutes in denaturation solution (0.5 M NaOH, 1.5 M NaCl), renatured for 5 minutes in renaturation solution (1 M Tris pH 7.5, 1.5 M NaCl), and washed briefly in 2xSSC (20x SSC: 3 M NaCl, 0.3 M Na-citrate). Filter membranes are dried and incubated at 80°C for 120 minutes to cross-link the phage DNA to the membranes.

The membranes are hybridized with a DNA probe prepared as described above. A DNA fragment (25 ng) is labeled with α-^{32}P-dCTP (NEN) using

Rediprime™ random priming (Amersham Pharmacia Biotech), according to manufacturers instructions. Labeled DNA is separated from unincorporated nucleotides by S200 spin columns (Amersham Pharmacia Biotech), denatured at 95°C for 5 minutes and kept on ice. The DNA-containing membranes (above) are pre-hybridized in 50 ml ExpressHyb™ (Clontech) solution at 68°C for 90 minutes. Subsequently, the labeled DNA probe is added to the hybridization solution, and the probe is left to hybridize to the membranes at 68°C for 70 minutes. The membranes are washed five times in 2x SSC, 0.1% SDS at 42°C for 5 minutes each, and finally washed 30 minutes in 0.1x SSC, 0.2% SDS. Filters are exposed to Kodak XAR™ film (Eastman Kodak Company, Rochester, N.Y., USA) with an intensifying screen at –80°C for 16 hours. One positive colony is isolated from the plates, and replated with about 1000 pfu on a 15 cm LB plate. Plating, plaque lift to filters and hybridization are performed as described above. About four positive phage plaques are isolated form this secondary screening.

cDNA containing plasmids (pBluescript SK-) are rescued from the isolated phages by in vivo excision by culturing XL-1 blue cells co-infected with the isolated phages and with the Excision helper phage, as described by manufacturer (Stratagene). XL-blue cells containing the plasmids are plated on LB plates and grown at 37°C for 16 hours. Colonies (18) from each plate are replated on LB plates and grown. One colony from each plate is stricken onto a nylon filter in an ordered array, and the filter is placed on a LB plate to raise the colonies. The filter is then hybridized with a labeled probe as described above. About three positive colonies are selected and grown up in LB medium. Plasmid DNA is isolated from the three clones by Qiagen Midi Kit™ (Qiagen) according to the manufacturer's instructions. The size of the insert is determined by digesting the plasmid with the restriction enzymes *NotI* and *SalI*, which establishes an insert size. The sequence of the entire insert is determined by automated sequencing on both strands of the plasmids.

nGPCR-1: PCR AND SUBCLONING

cDNAs were sequenced directly using an AB1377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI PRISM Ready Dye-Deoxy Terminator kit with Taq FS polymerase. Each ABI cycle sequencing reaction contained about 0.5μg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed

by 50 cycles: 98°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC® gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B table top centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 5μl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for three min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI373A files into the Sequencher program (Gene Codes, Ann Arbor, MI).

The PCR reaction was performed in 50μL samples containing 41.9μL H_2O, 5μL 10x Buffer containing 15 mM $MgCl_2$ (Boehringer Mannheim Expand High Fidelity PCR System), 0.5μL 10mM dNTP mix, 1.5μL human genomic DNA (Clontech #6550-1, 0.1μg/μL), 0.3μL primer VR1A (1μg/μL), 0.3μL primer VR1B (1μg/μL), and 0.5μL High Fidelity Taq polymerase (Boehringer Mannheim, 3.5U/μl). The primer sequences for and, respectively were:
5'TCAAAGCTTATGGAATCATCTTTCTCATTTGGAGTGATCCTTGCTGTC, (VR1A)(SEQ ID NO: 95) corresponding to the 5' end of the coding region and containing a *HindIII* restriction site, and:
5' TTCACTCGAGTTAGCCATCAAACTCTGAGCTGGAGATAGTGACGATGTG (VR1B)(SEQ ID NO: 96) corresponding to the 3' end of the coding region and containing an *Xho*I restriction site (Genosys). The PCR reaction was carried out using a GeneAmp PCR9700 thermocycler (Perkin Elmer Applied Biosystems) and started with 1 cycle of 94°C for 2 min followed by 5 cycles at 94°C for 30 sec, 60°C for 2 min, 72°C for 1.5 min, followed by 20 cycles at 94°C for 30 sec, 60°C for 30 sec, 72°C for 1.5 min.

The PCR reaction was loaded onto a 0.75% agarose gel. The DNA band was excised from the gel and the DNA eluted from the agarose using a QIAquick gel extraction kit (Qiagen). The eluted DNA was ethanol-precipitated and resuspended in 4μL H_2O for ligation. The ligation reaction consisted of 4μL of fresh ethanol-precipitated PCR product and 1μL of pCRII-TOPO vector (Invitrogen). The reaction was gently mixed and allowed to incubate for 5 min. at room temperature followed by

the addition of 1μL of 6x TOPO cloning stop solution and mixing for 10 sec. at room temperature. The sample was then placed on ice and 2μL was transformed in 50μL of One Shot cells (Invitrogen) and plated onto ampicillin plates. Four white colonies were chosen and the presence of an insert was verified by PCR in the following manner. Each colony was resuspended in 2 ml LB broth for 2 hrs. A 500μL aliquot was spun down in the microfuge, the supernatant discarded, and the pellet resuspended in 25μL of H_2O. A 16μL aliquot was removed and boiled for 5 min and the sample was placed on ice. The sample was microfuged briefly to pellet any bacterial debris and PCR was carried out with 15μL sample using primers VR1A and VR1B, described above.

Colonies from positive clones identified by PCR were used to inoculate a 4ml culture of LB medium containing 100 μg/ml ampicillin. Plasmid DNA was purified using the Wizard Plus Minipreps DNA purification system (Promega). Since the primers used to amplify the fragment of nGPCR-1 from genomic DNA were engineered to have *HindIII* and *XhoI* sites, the cDNA obtained from the minipreps was digested with these restriction enzymes. One clone was verified by gel electrophoresis to give a DNA band of the correct size. cDNA from this clone was then sequenced, yielding the sequence of SEQ ID NO: 73.

nGPCR-3: PCR AND SUBCLONING

First-strand cDNA synthesis was performed following the directions for 3'-RACE ready cDNA from the SMART™ RACE cDNA Amplification Kit (Clontech). First 3 μl of H_2O, 1 μl human whole brain poly A^+ RNA (1μg/μl) (Clontech, 6516-1) and 1 μl 3'-CDS primer were mixed together, incubated at 70°C for 2 minutes, then placed on ice for 2 minutes. Added to the tube was 2 μl 5X First-Strand buffer, 1 μl 20 mM DTT, 1 μl dNTP mix (10 mM) and 1 μl Superscript II RT (200 units/μl) (GIBCO/BRL). The tube was incubated at 42°C for 1.5 hours then the reaction was diluted with 250 μl of Tricine-EDTA buffer.

PCR was performed in a 50 µl reaction using components that come with the Advantage®-GC cDNA PCR Kit. The PCR reaction contained 22.4 µl H_2O, 10 µl 5X GC cDNA PCR Reaction buffer, 10 µl 5M GC Melt, 1µl 50X dNTP mix (10 mM each), 5 µl human brain cDNA, 0.3 µl of LW1649 (SEQ ID NO: 187)(1 µg/µl), 0.3 µl of LW1650 (SEQ ID NO: 188)(1 µg/µl), 1 µl 50X Advantage-GC cDNA polymerase mix. The PCR reaction was performed in a Perkin-Elmer 9600 GeneAmp PCR

System starting with 1 cycle of 94°C for 2 min then 8 cycles at 94°C for 15 sec, 72°C for 2 min (decreasing 1°C with each cycle), 72°C for 3 min, followed by 30 cycles of 94°C for 15 sec, 68°C for 3 min. The PCR reaction was loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed in a microcentrifuge. The eluted DNA was EtOH precipitated and resuspended in 4 H_2O for ligation. The PCR primer sequence for LW1649 was:

GCATAAGCTTGCCATGGGCCCCGGCGAGG (SEQ ID NO: 187)

and for LW1650 was:

GCATTCTAGACCTCAGTGTGTCTGCTGC (SEQ ID NO: 188). The underlined portion of the primers matches the 5' and 3' areas, respectively, of the coding region.

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4μl PCR product DNA, 1 μl Salt Solution and 1 μl pCRII-TOPO vector that was incubated for 5 minutes at room temperature and then placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 μl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The DNA subcloned into pCRII-TOPO was sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 μl of H_20, 8 μl of BigDye Terminator mix, 5 μl mini-prep DNA (0.1 μg/μl), and 1 μl primer (25 ng/μl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a Centriflex™ gel filtration cartridge, dried under vacuum, then dissolved in 16 μl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 95.

nGPCR-9: PCR AND SUBCLONING

The PCR reaction was performed in 50 μl containing 34.5 μl H_2O, 5 μl Buffer II (PE Applied Biosystems AmpliTaq Gold system), 6 μl 25 mM $MgCl_2$, 2 μl 10 mM dNTP mix, 1.5 μl human genomic DNA (Clontech #6550-1, 0.1 μg/μl), 0.3 μl primer VR9A (1 μg/μl), 0.3 μl primer VR9B (1 μg/μl), and 0.4 μl AmpliTaq Gold™ DNA Polymerase. The primer sequences for VR9A and VR9B were as follows:

VR9A 5'TTCAAAGCTTATGGAGTCGGGGCTGCTG 3' (SEQ ID NO: 101), corresponding to the 5' end of the coding region and containing a *HindIII* restriction site, and the reverse primer was:

VR9B 5' TTCACTCGAGTCAGTCTGCAGCCGGTTCTG 3', (SEQ ID NO: 102), corresponding to the 3' end of the coding region and containing an *Xho*I restriction site (Genosys). The PCR reaction was carried out using a GeneAmp PCR 9700 thermocycler (Perkin Elmer Applied Biosystems) and started with 1 cycle of 95°C for 10 min, then 10 cycles at 95°C for 30 sec, 72°C for 2 min decreasing 1°C each cycle, 72°C for 1 min, followed by 30 cycles at 95°C for 30 sec, 60°C for 30 sec, 72°C for 1 min. The PCR reaction was loaded on a 0.75% gel. The DNA band was excised from the gel and the DNA was eluted from the agarose using a QIAquick gel extraction kit (Qiagen). The eluted DNA was ethanol-precipitated and resuspended in 4 μl H_2O for ligation. The ligation reaction consisted of 4 μl of fresh ethanol-precipitated PCR product and 1 μl of pCRII-TOPO vector (Invitrogen). The reaction was gently mixed and allowed to incubate for 5 min at room temperature followed by the addition of 1 μl of 6x TOPO cloning stop solution and mixing for 10 sec at room temperature. The sample was then placed on ice and 2 μl was transformed in 50 μl of One Shot cells (Invitrogen) and plated onto ampicillin plates. Five white colonies were chosen and were used to inoculate a 4 ml culture of LB medium containing 100 μg/ml ampicillin. Plasmid DNA was purified using the Wizard Plus Minipreps DNA purification system (Promega). Since the primers used to PCR SEQ-9 from genomic DNA were engineered to have *HindIII* and *XhoI* sites, the cDNA obtained from the minipreps was digested with these restriction enzymes. One clone was verified by gel electrophoresis to give a DNA band of the correct size. cDNA from this clone was then submitted for sequencing. One mutation was found (bp 621 T→G) and repaired as described as below.

The mutation in the identified clone was repaired using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The PCR reaction contained 39.3 μl H_2O, 5 μl 10x reaction buffer, 50 ng mini-prep cDNA, 1.25 μl primer VR9E (100 ng/μl), 1.25 μl primer VR9F (100 ng/μl), 1 μl 20 mM dNTP mix, 1 μl Pfu DNA polymerase. The cycle conditions were 95°C for 30 sec, then 12 cycles at 95°C for 30 sec, 55°C for 1 min, 68°C for 10 min. One μl of DpnI was added and the tube incubated at 37°C for 1 hr. One μl of the DpnI-treated DNA was transformed into 50 μl Epicurian coli XL1-Blue supercompetent cells and the entire insert was re-sequenced. The primer sequences used were:

VR9E: 5' GCATCCTGGCCGC<u>T</u>ATCTGTGCACTCTACG 3' (SEQ ID NO: 103) and

VR9F: 5' CGTAGAGTGCACAGAT<u>A</u>GCGGCCAGGATGC 3' (SEQ ID NO: 104) where the base underlined was the base being corrected.

The clone described above was sequenced directly using an ABI377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI BigDye™ Terminator Cycle Sequencing Ready Reaction kit with Taq FS™ polymerase. Each ABI cycle sequencing reaction contained 0.5 μg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 96°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC (R) gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 3 μl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 3.5 min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 77.

nGPCR-11: PCR AND SUBCLONING

PCR was performed in a 50 µl reaction containing 32 µl H_2O, 5 µl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 µl 15 mM $MgSO_4$, 2 µl 10 mM dNTP, 5 µl human genomic DNA (0.3µg/µl)(Clontech), 0.3 µl of LW1564 (1 µg/µl), 0.3 µl of LW1565 (1 µg/µl), 0.4 µl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was performed in a GeneAmp 9600 PCR thermocycler (PE Applied Biosystems) starting with 1 cycle of 94°C for 2 min followed by 17 cycles at 94°C for 30 sec, 72°C for 2 min decreasing 1°C each cycle, 68°C for 2 min, then 25 cycles of 94°C for 30 sec, 55°C for 30 sec, 68°C for 2 min. The PCR reaction was loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed in a microcentrifuge. The eluted DNA was EtOH precipitated and resuspended in 4µl H_2O for ligation. The forward PCR primer sequence was:

LW1564: GCATAAGCTTCCATGTACAACGGGTCGTGCTGC (SEQ ID NO: 107), and the reverse PCR primer was:

LW1565: GCATTCTAGATCAGTGCCACTCAACAATGTGGG (SEQ ID NO: 108).

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4 µl PCR product DNA and 1 µl pCRII-TOPO vector that was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 Tl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The DNA subcloned into pCRII was sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 µl of

H_20, 8 µl of BigDye Terminator mix, 5 µl mini-prep DNA (0.1 µg/µl), and 1 µl primer (25 ng/µl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a Centriflex™ gel filtration cartridge, dried under vacuum, then dissolved in 16 µl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 79.

nGPCR-16: PCR AND SUBCLONING

PCR was performed in a 50 µl reaction containing 32 µl H_2O, 5 µl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 µl 15 mM $MgS0_4$, 2 µl 10 mM dNTP, 5 µl 2445704H1 DNA (0.17 Tg/Tl), 0.3 µl of LW1587 (1 µg/µl), 0.3 µl of LW1588 (1 µg/µl), 0.4 µl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was performed on a Robocycler thermocycler (Stratagene) starting with 1 cycle of 94°C for 2 min followed by 15 cycles of 94°C for 30 sec, 55°C for 1.3 min, 68°C for 2 min. The PCR reaction was loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed in a microcentrifuge. The eluted DNA was EtOH precipitated and resuspended in 12µl H_2O for ligation. The PCR primer sequence for the forward primer was:

LW1587: GATCAAGCTTATGACAGGTGACTTCCCAAGTATGC (SEQ ID NO: 111), and the sequence for the reverse primer was:

LW1588: GATCCTCGAGGCTAACGGCACAAAACACAATTCC (SEQ ID NO: 112).

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4µl PCR product DNA and 1 µl pCRII-TOPO vector that was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 µl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium.

Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The DNA subcloned into pCRII was sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 µl of H_20, 8 µl of BigDye Terminator mix, 5 µl mini-prep DNA (0.1 µg/µl), and 1 µl primer (25 ng/µl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a Centriflex™ gel filtration cartridge, dried under vacuum, then dissolved in 16 µl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 81.

nGPCR-40: PCR AND SUBCLONING

PCR was performed in a 50 µl reaction containing utilizing Herculase DNA Polymerase blend (Stratagene), using the buffer recommendations provided by the manufacturer, 200 ng each of primers PSK 18 and 19 (SEQ ID NOS: 115 and 116), 150 ng of human genomic DNA (Clontech), and 2% DMSO. The PCR reaction was performed on a Robocycler thermocycler (Stratagene) starting with 1 cycle of 94°C for 2 min followed by 35 cycles of 94°C for 30 sec, 65°C for 30 sec, 72°C for 2 min. The PCR reaction was purified using the QiaQuick PCR Purification Kit (Qiagen), and then eluted in TE. The PCR primer sequences were:

PSK 18 GATC GAATTCGCAGGAGCAATG AAAATCAGGAAC (SEQ ID NO: 115), and:

PSK19: GATCGAATTC**TTA**TATATGTTCAGAAAACAAATTCATGG (SEQ ID NO: 116)). The underlined portion of the primer matches the 5' and 3' areas, respectively, of a portion of the 5' untranslated region and coding region. Initiation and termination codons are shown above in bold.

The PCR product was ligated into the pCR-BluntII-TOPO vector (Invitrogen) using the Zero Blunt Topo PCR TA cloning kit as follow: 3µl PCR product DNA, 1 µl pCRII-TOPO vector, and 1 µl TOPOII salt solution (1.2M NaCl, 0.06M $MgCl_2$). The mixture was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added, and then the

reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 μl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates supplemented with Xgal and IPTG. Single colonies were screened by PCR for the presence of the insert, and a plasmid DNA from colony 58 was purified using a Qiagen Endo-Free plasmid purification kit.

nGPCR-40 was sequenced directly using an ABI377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI BigDye™ Terminator Cycle Sequencing Ready Reaction kit with Taq FS™ polymerase. Each ABI cycle sequencing reaction contained about 0.5 µg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 96°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC® gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 3 µl of DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 3.5 min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the Sequencher program (Gene Codes, Ann Arbor, MI), which yielded a sequence identical to SEQ ID NO:83 with the exception that the nucleotide at position 10 was identified as an "A" which incorrectly indicated the presence of an initiation codon at that position. Subsequent analysis of genomic DNA samples indicated that this position was incorrectly assigned and that the correct nucleotide at that position was a "C". The sequence reported at SEQ ID NO. 83 correctly identifies the nucleotide at position 10 and indicates that the first initiation codon occurs at position 88-90.

nGPCR-54: PCR AND SUBCLONING

Two microliters of a human genomic library (~10^8 PFU/ml) (Clontech) was added to 6 ml of an overnight culture of K802 cells (Clontech), then distributed as 250 µl aliquots into each of 24 tubes. The tubes were incubated at 37°C for 15 min. Seven milliliters of 0.8% agarose was added to each tube, mixed, then poured onto LB agar + 10 mM $MgSO_4$ plates and incubated overnight at 37°C. To each plate 5 ml of SM (0.1M NaCl, 8.1 mM $MgSO_4$-$7H_2O$, 50mM Tris-Cl (pH 7.5), 0.0001% gelatin) phage buffer was added and the top agarose was removed with a microscope slide and placed in a 50 ml centrifuge tube. A drop of chloroform was added and the tube was place in a 37 °C shaker for 15 min, then centrifuged for 20 min at 4000 RPM (Sorvall RT6000 table top centrifuge) and the supernatant stored at 4°C as a stock solution.

Two µl of phage from each tube was heated to 99°C for 4 min then cooled to 10°C. Added to the phage was a PCR mix containing 8.8 µl H_2O, 4 µl 5X Rapid-Load Buffer (Origene), 2 µl 10xPCR buffer II (Perkin-Elmer), 2 µl 25 mM $MgCl_2$, 0.8 µl 10 mM dNTP, 0.12 µl LW1634 (1 µg/µl)(SEQ ID NO: 119), 0.12 µl LW1635 (1 µg/µl)(SEQ ID NO: 120), 0.2 µl AmpliTaq Gold polymerase (Perkin Elmer). The PCR reaction involved 1 cycle at 95°C for 10 min followed by 35 cycles at 95°C for 45 sec, 53.5°C for 2 min, 72°C for 45 sec. The reaction was loaded onto a 2 % agarose gel. From the tube that gave a PCR product of the correct size, 10 µl was used to set up five 1:10 dilutions that were plated onto LB agar + 10 mM $MgSO_4$ plates and incubated overnight. A BA85 nitrocellulose filter (Schleicher & Schuell) was placed on top of each plate for 1 hour. The filter was removed, placed phage side up in a petri dish, and covered with 4 ml of SM for 15 min to elute the phage. One milliliter of SM was removed from each plate and used to set up a PCR reaction as above. The plate of the lowest dilution to give a PCR product was subdivided, filter-lifted and the PCR reaction was repeated. The series of dilutions and subdividing of the plate was continued until a single plaque was isolated that gave a positive PCR band. Once a single plaque was isolated, 10 µl phage supernatant was added to 100 µl SM and 200 µl of K802 cells per plate with a total of 8 plates set up. The plates were incubated overnight at 37°C. The top agarose was removed by adding 8 ml of SM then scrapping off the agarose with a microscope slide and collected in a centrifuge tube. To the tube, 3 drops of chloroform was added, vortexed, incubated at 37°C for 15 min then centrifuged for 20 min at 4000 RPM (Sorvall RT6000 table top

centrifuge) to recover the phage, which was used to isolate genomic phage DNA using the Qiagen Lambda Midi Kit. The sequence for primer LW1634 was:

CTGAAAGTTGTCGCTGACC (SEQ ID NO: 119), and for primer LW1635 was:

CGATTATCCACACTTTGACCC (SEQ ID NO: 120).

The PCR reaction for the coding region was performed in a 50 μl reaction containing 33 μl H_2O, 5 μl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 μl 15 mM $MgSO_4$, 2 μl 10 mM dNTP, 4 μl genomic phage DNA (0.25 μg/μl), 0.3 μl LW1698 (1 μg/μl)(SEQ ID NO: 121), 0.3 μl LW1699 (1 μg/μl)(SEQ ID NO: 122), 0.4 μl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was started with 1 cycle of 94°C for 2 min followed by 30 cycles at 94°C for 30 sec, 55°C for 30 sec., 68°C for 2 min. The PCR reaction was loaded onto a 2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA was EtOH precipitated and resuspended in 8μl H_2O. The PCR primer sequence for primer LW1698 was:

GCATACC<u>ATGAATGAGCCACTAGAC</u> (SEQ ID NO: 121), and for primer LW1699 was:

GCATCTCGAG<u>TCAAGGGTTGTTTGAGTAAC</u> (SEQ ID NO: 122). The underlined portion of the primer matches the 5' and 3' areas, respectively, of the coding region of nGPCR-54.

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4μl PCR product DNA, 1 μl of salt solution and 1 μl pCRII-TOPO vector that was incubated for 5 minutes at room temperature then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 μl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

nGPCR-54 genomic phage DNA was sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary

electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. The cycle-sequencing reaction contained 14 μl of H_20, 16 μl of BigDye Terminator mix, 7 μl genomic phage DNA (0.1 μg/μl), and 3 μl primer (25 ng/μl). The reaction was performed in a Perkin-Elmer 9600 thermocycler at 95°C for 5 min, followed by 99 cycles of 95°C for 30 sec, 55°C for 20 sec, and 60°C for 4 min. The product was purified using a Centriflex™ gel filtration cartridges, dried under vacuum, then dissolved in 16 μl of Template Suppression Reagent. The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer.

The DNA subcloned into pCRII was sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 μl of H_20, 8 μl of BigDye Terminator mix, 5 μl mini-prep DNA (0.1 μg/μl), and 1 μl primer (25 ng/μl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a Centriflex™ gel filtration cartridge, dried under vacuum, then dissolved in 16 μl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 85.

nGPCR-56: PCR AND SUBCLONING

The PCR reaction for the coding region of nGPCR-56 used components that come with PLATINUM® *Pfx* DNA Polymerase (GibcoBRL) containing 35.5 μl H_2O, 5 μl 10X Pfx Amplification buffer, 1.5 μl 50mM $MgS0_4$, 2 μl 10 mM dNTP, 5 μl human genomic DNA (0.3μg/μl)(Clontech), 0.3 μl of LW1603 (1 μg/μl)(SEQ ID NO: 152), 0.3 μl of LW1604 (1 μg/μl)(SEQ ID NO: 153), 0.4 μl PLATINUM® *Pfx* DNA Polymerase (2.5 U/Tl). The PCR reaction was performed in a Robocycler Gradient 96 (Stratagene) starting with 1 cycle of 94°C for 5 min followed by 30 cycles at 94°C for 40 sec, 55°C for 2 min, 68°C for 3 min. Following the final cycle, 0.5 μl of AmpliTaq DNA Polymerase (5 U/μl) was added and the tube was incubated at 72°C for 5 min. The sequence of LW1603 is:

GATCAAGCTTGGAATGATGCCCTTTTGCCAC (SEQ ID NO: 152), and for LW1604 is:

GATCCTCGAGCATCATTCAAAGTAGGTGG. (SEQ ID NO: 153). The underlined portion of the primer matches the 5' and 3' areas, respectively, of a portion of the coding region of nGPCR-56.

The PCR reaction for the coding region was performed in a 50 μl reaction containing 32 μl H_2O, 5 μl 10X TT buffer (140 mM Ammonium Sulfate, 0.1 % gelatin, 0.6 M Tris-tricine pH 8.4), 5 μl 15 mM $MgSO_4$, 2 μl 10 mM dNTP, 5 μl human genomic DNA (0.3μg/μl)(Clontech), 0.3 μl LW1603 (1 μg/μl)(SEQ ID NO: 152), 0.3 μl LW1696 (1 μg/μl)(SEQ ID NO: 154), 0.4 μl High Fidelity Taq polymerase (Boehringer Mannheim). The PCR reaction was started with 1 cycle of 94°C for 2 min followed by 25 cycles at 94°C for 40 sec, 55°C for 60 sec., 68°C for 2 min. The PCR reaction was loaded onto a 2 % agarose gel. The DNA band was excised from the gel, placed in GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA was EtOH precipitated and resuspended in 12μl H_2O for ligation. The PCR primer sequence for LW1603 is:

GATCAAGCTTGGAATGATGCCCTTTTGCCAC (SEQ ID NO: 152), and LW1696:

GATCCTCGAGCTATGAACTCAATTCCAAAAATAATTTACACC (SEQ ID NO: 154). The underlined portion of the primer matches the 5' and 3' areas, respectively, of a portion of the coding region.

The ligation reaction used solutions from the TOPO TA Cloning Kit (Invitrogen) which consisted of 4μl PCR product DNA, 1 μl of salt solution and 1 μl pCRII-TOPO vector that was incubated for 5 minutes at room temperature then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 μl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates. A single colony containing an insert was used to inoculate a 5 ml culture of LB medium. Plasmid DNA was purified using a Concert Rapid Plasmid Miniprep System (GibcoBRL) and then sequenced.

The mutation in nGPCR-56 was repaired using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The PCR reaction contained 40 μl H2O, 5 μl 10x

Reaction buffer, 1 µl mini-prep DNA, 1 µl LW1700 (125 ng/µl) (SEQ ID NO: 155), 1µl LW1701 (125 ng/µl) (SEQ ID NO: 156), 1µl 10 mM dNTP, 1 µl Pfu DNA polymerase. The cycle conditions were 95°C for 30 sec then 14 cycles at 95°C for 30 sec, 55°C for 1 min, 68°C for 12 min. The tube was placed on ice for 2 min, then 1 µl of *DpnI* was added and the tube incubated at 37°C for one hour. One microliter of the *DpnI*-treated DNA was transformed into Epicurian coli XL1-Blue supercompetent cells and the entire insert was re-sequenced. The primer sequences are: GCTACTTGAACTCTACATTTAATCCAATGGTTTATGCATTTTTCTATCC (LW1700)(SEQ ID NO: 155), and: GGATAGAAAAATGCATAAACCATTGGATTAAATGTAGAGTTCAAGTAGC (LW1701)(SEQ ID NO: 156).

The DNA subcloned into pCRII was sequenced using the ABI PRISM™ 310 Genetic Analyzer (PE Applied Biosystems) which uses advanced capillary electrophoresis technology and the ABI PRISM™ BigDye™ Terminator Cycle Sequencing Ready Reaction Kit. Each cycle-sequencing reaction contained 6 µl of H_20, 8 µl of BigDye Terminator mix, 5 µl mini-prep DNA (0.1 µg/µl), and 1 µl primer (25 ng/µl) and was performed in a Perkin-Elmer 9600 thermocycler with 25 cycles of 96°C for 10 sec, 50°C for 10 sec, and 60°C for 4 min. The product was purified using a Centriflex™ gel filtration cartridge, dried under vacuum, then dissolved in 16 µl of Template Suppression Reagent (PE Applied Biosystems). The samples were heated at 95°C for 5 min then placed in the 310 Genetic Analyzer, yielding the sequence of SEQ ID NO: 89.

nGPCR-58: PCR AND SUBCLONING

Isolation of a clone for nGPCR-58 from genomic DNA was performed by PCR in a 50 µl reaction containing Herculase DNA Polymerse blend (Stratagene), with buffer recommendations as supplied by the manufacturer, 200 ng each primers PSK14 (SEQ ID NO: 157) and PSK15 (SEQ ID NO: 158), 150 ng of human genomic DNA (Clontech) and 6% DMSO. The PCR reaction was performed on a Robocycler thermocycler (Stratagene) starting with 1 cycle of 94°C for 2 min followed by 35 cycles of 94°C for 30 sec, 65°C for 30 sec, 72°C for 2 min. The PCR reaction was purified by the QiaQuick PCR Purification Kit (Qiagen) and eluted in TE. The PCR primer sequences were:

PSK14: 5'GATCGAATTC**<u>ATG</u>**<u>GACACTACCATGGAAGCTGACC</u> (SEQ ID NO: 157), and:

PSK15: 5'GATCCTCGAG**<u>TCA</u>**<u>CGTGGGGCCTGCGCCCGG</u> (SEQ ID NO: 158).

The underlined portion of the primers match the 5' and 3' areas, respectively, of a portion of the 5' untranslated region and coding region. Translation initiation and termination codons are shown above in bold.

The blunt ended PCR product was prepared for cloning by the addition of a single base "A" residue by AmpliTaq Gold (Perkin Elmer) in a reaction with 1X PCR Buffer II, 1 mM $MgCl_2$, 200uM each dATP, dGTP, dCTP, and dTTP. The reaction was incubated at 94°C for 10 minutes followed by 72°C for 10 minutes. The products were cloned into the pCRII-TOPO vector (Invitrogen) using the TOPO TA cloning kit as follows: 3μl PCR product DNA , 1 μl pCRII-TOPO vector, and 1 μl TOPOII salt solution (1.2M NaCl, 0.06M $MgCl_2$) was incubated for 5 minutes at room temperature. To the ligation reaction one microliter of 6X TOPO Cloning Stop Solution was added then the reaction was placed on ice. Two microliters of the ligation reaction was transformed in One-Shot TOP10 cells (Invitrogen), and placed on ice for 30 minutes. The cells were heat-shocked for 30 seconds at 42°C, placed on ice for two minutes, 250 μl of SOC was added, then incubated at 37°C with shaking for one hour and then plated onto ampicillin plates supplemented with X-gal and IPTG. Single colonies were screened by PCR for the presence of the insert, and a plasmid DNA from colony 58-6 was purified using a Qiagen Endo-Free plasmid purification kit and deposited as nGPCR-58.

nGPCR-58 was sequenced directly using an ABI377 fluorescence-based sequencer (Perkin Elmer/Applied Biosystems Division, PE/ABD, Foster City, CA) and the ABI BigDye™ Terminator Cycle Sequencing Ready Reaction kit with Taq FS™ polymerase. Each ABI cycle sequencing reaction contained about 0.5 μg of plasmid DNA. Cycle-sequencing was performed using an initial denaturation at 98°C for 1 min, followed by 50 cycles: 96°C for 30 sec, annealing at 50°C for 30 sec, and extension at 60°C for 4 min. Temperature cycles and times were controlled by a Perkin-Elmer 9600 thermocycler. Extension products were purified using AGTC (R) gel filtration block (Edge BiosSystems, Gaithersburg, MD). Each reaction product was loaded by pipette onto the column, which was then centrifuged in a swinging

bucket centrifuge (Sorvall model RT6000B tabletop centrifuge) at 1500 x g for 4 min at room temperature. Column-purified samples were dried under vacuum for about 40 min and then dissolved in 3 μl of a DNA loading solution (83% deionized formamide, 8.3 mM EDTA, and 1.6 mg/ml Blue Dextran). The samples were then heated to 90°C for 3.5 min and loaded into the gel sample wells for sequence analysis by the ABI377 sequencer. Sequence analysis was performed by importing ABI377 files into the Sequencer program (Gene Codes, Ann Arbor, MI), yielding the sequence of SEQ ID NO: 93.

EXAMPLE 3: HYBRIDIZATION ANALYSIS TO DEMONSTRATE nGPCR-X EXPRESSION IN BRAIN

The expression of nGPCR-x in mammals, such as the rat, may be investigated by *in situ* hybridization histochemistry. To investigate expression in the brain, for example, coronal and sagittal rat brain cryosections (20 μm thick) are prepared using a Reichert-Jung cryostat. Individual sections are thaw-mounted onto silanized, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. Sections are processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold phosphate-buffered saline (PBS), acetylated using acetic anhydride in triethanolamine buffer, and dehydrated through a series of alcohol washes in 70%, 95%, and 100% alcohol at room temperature. Subsequently, sections are delipidated in chloroform, followed by rehydration through successive exposure to 100% and 95% alcohol at room temperature. Microscope slides containing processed cryosections are allowed to air dry prior to hybridization. Other tissues may be assayed in a similar fashion.

A nGPCR-x-specific probe is generated using PCR. Following PCR amplification, the fragment is digested with restriction enzymes and cloned into pBluescript II cleaved with the same enzymes. For production of a probe specific for the sense strand of nGPCR-x, the nGPCR-x clone in pBluescript II is linearized with a suitable restriction enzyme, which provides a substrate for labeled run-off transcripts (*i.e.*, cRNA riboprobes) using the vector-borne T7 promoter and commercially available T7 RNA polymerase. A probe specific for the antisense strand of nGPCR-x is also readily prepared using the nGPCR-x clone in pBluescript II by cleaving the recombinant plasmid with a suitable restriction enzyme to generate a linearized substrate for the production of labeled run-off cRNA transcripts using the T3

promoter and cognate polymerase. The riboprobes are labeled with [^{35}S]-UTP to yield a specific activity of about 0.40 x 10^6 cpm/pmol for antisense riboprobes and about 0.65 x 10^6 cpm/pmol for sense-strand riboprobes. Each riboprobe is subsequently denatured and added (2 pmol/ml) to hybridization buffer which contained 50% formamide, 10% dextran, 0.3 M NaCl, 10 mM Tris (pH 8.0), 1 mM EDTA, 1X Denhardt's Solution, and 10 mM dithiothreitol. Microscope slides containing sequential brain cryosections are independently exposed to 45 μl of hybridization solution per slide and silanized cover slips are placed over the sections being exposed to hybridization solution. Sections are incubated overnight (15-18 hours) at 52°C to allow hybridization to occur. Equivalent series of cryosections are exposed to sense or antisense nGPCR-40-specific cRNA riboprobes.

Following the hybridization period, coverslips are washed off the slides in 1X SSC, followed by RNase A treatment involving the exposure of slides to 20 μg/ml RNase A in a buffer containing 10 mM Tris-HCl (pH 7.4), 0.5 M EDTA, and 0.5 M NaCl for 45 minutes at 37°C. The cryosections are then subjected to three high-stringency washes in 0.1 X SSC at 52°C for 20 minutes each. Following the series of washes, cryosections are dehydrated by consecutive exposure to 70%, 95%, and 100% ammonium acetate in alcohol, followed by air drying and exposure to Kodak BioMax™ MR-1 film. After 13 days of exposure, the film is developed. Based on these results, slides containing tissue that hybridized, as shown by film autoradiograms, are coated with Kodak NTB-2 nuclear track emulsion and the slides are stored in the dark for 32 days. The slides are then developed and counterstained with hematoxylin. Emulsion-coated sections are analyzed microscopically to determine the specificity of labeling. The signal is determined to be specific if autoradiographic grains (generated by antisense probe hybridization) are clearly associated with cresyl violate-stained cell bodies. Autoradiographic grains found between cell bodies indicates non-specific binding of the probe.

Expression of nGPCR-x in the brain provides an indication that modulators of nGPCR-x activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-x may have utility include depression,

anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-x modulators, including nGPCR-x ligands and anti-nGPCR-x antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

EXAMPLE 4: TISSUE EXPRESSION PROFILING

Tissue specific expression of the cDNAs encoding nGPCR-1, nGPCR-3, nGPCR-9, nGPCR-11, nGPCR-16, nGPCR-40, nGPCR-54, nGPCR-56, and nGPCR-58 was detected using a PCR-based system. Tissue specific expression of cDNAs encoding nGPCR-x may be accomplished using similar methods.

Primers were synthesized by Genosys Corp., The Woodlands, TX. PCR reactions were assembled using the components of the Expand Hi-Fi PCR System™ (Roche Molecular Biochemicals, Indianapolis, IN).

nGPCR-1

The RapidScan™ Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array may include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. Human brain regions in the array may include: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of the nGPCR-1 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 212bp fragment in the presence of the appropriate cDNA. The forward primer was:

GCTCAACCCACTCATCTATGCC (SEQ ID NO: 97), and the reverse primer was:

AAACTTCTCTGCCCTTACCGTC (SEQ ID NO: 98)

The PCR reaction mixture was added to each well of the PCR plate. The plate was placed in a GeneAmp PCR9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The plate was then exposed to the following cycling parameters: Pre-soak 94°C for 3 min; denaturation at 94°C for 30 seconds; annealing at primer T_m for

45 seconds; extension 72°C for 2 minutes; for 35 cycles. PCR products were then separated and analyzed by electrophoresis on a 1.5-% agarose gel.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction is within the linear range and, hence, facilitated the semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

Expression of nGPCR-1 was found to be highest in the testis, adrenal gland and heart. Significant levels of expression were also found in the brain, kidney, spleen ovary, prostate, muscle, PBL, stomach and bone marrow. Within the brain, expression levels were highest in the cerebellum, amygdala, thalamus and spinal cord, with significant levels of expression in the frontal lobe, hippocampus, substantia nigra, hypothalamus and pons.

Expression of nGPCR-1 in the brain provided an indication that modulators of nGPCR-1 activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-1 may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthénia, neuropathy, neuroses, and the like. Use of nGPCR-1 modulators, including nGPCR-1 ligands and anti-nGPCR-1 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-3

Tissue specific expression of the cDNA encoding nGPCR-3 was detected using a PCR-based method. Multiple Choice™ first strand cDNAs (OriGene Technologies, Rockville, MD) from 6 human tissues were serially diluted over a 3-log range and arrayed into a multi-well PCR plate. This array was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed included: brain, heart, kidney, peripheral blood leukocytes, lung and testis. PCR primers were designed based on the available sequence of the putative GPCR. The sequence of the forward primer used was:

5'TGCTGCTTTGTTGCGCCTAC3' (SEQ ID NO: 189), corresponding to base pairs 77 through 96 of the predicted coding sequence of nGPCR-3. The sequence of the reverse primer used was:

5'TTGGACGCCAGGAAGGTG3' (SEQ ID NO: 190), corresponding to base pairs 258 through 285 of the predicted coding sequence of nGPCR-3. This primer set primes the synthesis of a 298 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid Scan™ Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 3 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord. Primers were synthesized by Genosys Corp., The Woodlands, TX. PCR reactions were assembled using the components of the Expand Hi-Fi PCR System™ (Roche Molecular Biochemicals, Indianapolis, IN). Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94°C for 45 sec.), (53°C for 2 min.), and (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

The results indicated that nGPCR-3 was expressed in the brain, heart, kidney, peripheral blood lymphocytes, lung, and testis. In the brain, nGPCR-3 was expressed in frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla, as well as in the spinal cord.

nGPCR-9

The RapidScan™ Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver.

The forward primer used was to detect expression of nGPCR-9 was:

5' AACCCCATCATCTACACGC 3'(SEQ ID NO: 105), and, the reverse primer was:

5' TGCCTGTGGAGCCGCTGG 3'(SEQ ID NO: 106). This primer set will prime the synthesis of a 238 base pair fragment in the presence of the appropriate cDNA.

For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid Scan™ Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2-log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°Cfor 3 min.) followed by 35 cycles of [(94°Cfor 45 sec.) (52°C for 2 min.) (72°Cfor 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel and stained with ethidium bromide.

nGPCR-9 was expressed in the brain, peripheral blood leukocytes, heart, kidney, adrenal gland, spleen, pancreas, liver, lung, skin, bone marrow, testis, placenta, salivary gland, uterus, small intestine, muscle, stomach, and fetal liver. Within the brain, nGPCR-9 was expressed in all areas examined including the frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-9 in the brain provided an indication that modulators of nGPCR-9 activity have utility for treating disorders, including but not limited to, schizophrenia, affective disorders, movement disorders, metabolic disorders, inflammatory disorders, cancers, ADHD/ADD (i.e., Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Use of nGPCR-9 modulators, including nGPCR-9 ligands and anti-nGPCR-9 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-11

The RapidScan™ Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array included, *inter alia*: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal

gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. Human brain regions in the array included, *inter alia*: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-11 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 206bp fragment in the presence of the appropriate cDNA. The forward primer used to detect expression of nGPCR-11 was:

5'-GAAGCCCAGCACTGTTTACC-3' (SEQ ID NO: 109), and the reverse primer was:

5'-TGAAATACCTGTCCGCAGCC-3 (SEQ ID NO: 110).

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (PE Applied Biosystems). The following cycling program was executed: Pre-soak 94°C for 3 min; denaturation at 94°C for 30 seconds; annealing at primer T_m for 45 seconds; extension at 72°C for 2 minutes; for 35 cycles. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction was within the linear range and, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-11 was expressed in the thyroid gland, brain, heart, kidney, adrenal gland, spleen, liver, ovary, muscle, testis, salivary gland, colon, prostate, small intestine, skin stomach, bone marrow, fetal brain and placenta. Within the brain, nGPCR-11 was expressed in the temporal lobe, amygdala, substantia nigra, pons, spinal cord, frontal lobe, and cerebellum.

Expression of the nGPCR-11 in the brain provided an indication that modulators of nGPCR-11 activity have utility for treating disorders, including but not limited to, schizophrenia, affective disorders, metabolic disorders, inflammatory disorders, cancers, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-11 may have utility include depression, anxiety, bipolar

disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-11 modulators, including nGPCR-11 ligands and anti-nGPCR-11 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

Expression of nGPCR-11 in the thyroid gland, indicates that agonists or antagonists could be of use in the treatment of thyroid dysfunction such as thyreotoxicosis and myxoedema. They could also be of use in the stimulation of thyroid hormone release leading to overall increase in metabolic rate and weight reduction. The expression of nGPCR-11 in liver and muscle indicate a use for agonists or antagonists in regulation of glucose metabolism applicable in diabetes type II.

nGCPR-16

The RapidScan™ Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array included, *inter alia*: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. Human brain regions in the array included, *inter alia*: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-16 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 205bp fragment in the presence of the appropriate cDNA. The forward primer used to detect expression of nGPCR-16 was:

5' CAGCCCAAACATCCAAGTC 3'. (SEQ ID NO: 113). The reverse primer used to detect expression of nGPCR-16 was:

5' ACCCCACTTAATCAGCCTC 3'(SEQ ID NO: 114).

For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid Scan™ Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94° for 3min.) followed by 35 cycles of [(94°C for 45 sec.) (53°C for 2 min.) (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel, and stained with ethidium bromide.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction was within the linear range and, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-16 was expressed in the ovary, lung, prostate, bone marrow, salivary gland, heart, adrenal gland, spleen, liver, small intestine, skin, muscle, peripheral blood leukocytes, testis, placenta, fetal liver, brain, thyroid gland, kidney, pancreas, colon, uterus, and stomach.. Within the brain, nGPCR-16 was expressed in all areas examined including the frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of nGPCR-16 in the brain provides an indication that modulators of nGPCR-16 activity have utility for treating neurological disorders, including but not limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), and neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, and senile dementia. Some other diseases for which modulators of nGPCR-16 may have utility include depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, and the like. Use of nGPCR-16 modulators, including nGPCR-16 ligands and anti-nGPCR-16 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-40

The RapidScan™ Gene Expression Panel (OriGene Technologies, Rockville, MD) was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid,

adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. The forward primer used was:

5'ACAGCCCCAAAGCCAAACAC3', (SEQ ID NO: 117), and the reverse primer was:

5'CCGCAGGAGCAATGAAAATCAG3', (SEQ ID NO: 118). This primer set primed the synthesis of a 220 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain RapidScan™ Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94° for 45 sec.) (54°C for 2 min.) (72° for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

The dilution range of cDNA deposited on the plates ensured that the amplification reaction was within the linear range and, hence, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-40 was expressed in the brain, peripheral blood lymphocytes, pancreas, ovary, uterus, testis, salivary gland, kidney, adrenal gland, liver, bone marrow, prostate, fetal liver, colon, muscle, and fetal brain, may be found in many other tissues, including, but not limited to, lung, small intestine, fetal brain cord, and bone. Within the brain, nGPCR-40 was expressed in the frontal lobe, hypothalamus, pons, cerebellum, caudate nucleus, and medulla.

Expression of nGPCR-40 in the brain provides an indication that modulators of nGPCR-40 activity have utility for treating neurological disorders, including but not limited to, movement disorders, affective disorders, metabolic disorders, inflammatory disorders and cancers. Use of nGPCR-40 modulators, including nGPCR-40 ligands and anti-nGPCR-40 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-54

Multiple Choice™ first strand cDNAs (OriGene Technologies, Rockville, MD) from 12 human tissues were serially diluted over a 3-log range and arrayed into a multi-well PCR plate. Human tissues arrayed include: brain, heart, kidney, peripheral blood leukocytes, liver, lung, muscle, ovary, prostate, small intestine, spleen and testis. PCR primers were designed based on the sequence of nGPCR-54 provided herein. The forward primer used was:

5'CTGTCTCTCTGTCCTCTTCC3',(SEQ ID NO: 123). The reverse primer used was:

5'GCACCGATCTTCATTGAATTTC3',(SEQ ID NO: 124). This primer set primes the synthesis of a 145 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid Scan™ Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 3 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94°C for 45 sec.) (52.5°C for 2 min.) (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

nGPCR-54 was expressed in the brain, kidney, lung, muscle, testis, heart, liver, ovary, prostate, small intestine, spleen, and peripheral blood leukocytes. Within the brain, nGPCR-54 was expressed in the cerebellum, hippocampus, substantia nigra, thalamus, hypothalamus, pons, frontal lobe, temporal lobe, caudate nucleus, medulla, spinal cord, and amygdala.

Expression of the nGPCR-54 in the brain provides an indication that modulators of nGPCR-54 activity have utility for treating neurological disorders, including but not limited to, movement disorders, affective disorders, metabolic disorders, inflammatory disorders and cancers. Use of nGPCR-54 modulators,

including nGPCR-54 ligands and anti-nGPCR-54 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-56

The RapidScan™ Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues arrayed include: brain, heart, kidney, spleen, liver, colon, lung, small intestine, muscle, stomach, testis, placenta, salivary gland, thyroid, adrenal gland, pancreas, ovary, uterus, prostate, skin, PBL, bone marrow, fetal brain, fetal liver. The forward primer used was:

5' ACTTCAAACAACTTCATACCCC 3' (SEQ ID NO: 125), and the reverse primer used was:

5'ACACACAGCATAGTAGCG 3' (SEQ ID NO: 126). This primer set will prime the synthesis of a 231 base pair fragment in the presence of the appropriate cDNA. For detection of expression within brain regions, the same primer set was used with the Human Brain Rapid Scan™ Panel (OriGene Technologies, Rockville, MD). This panel represents serial dilutions over a 2 log range of first strand cDNA from the following brain regions arrayed in a 96 well format: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Twenty-five microliters of the PCR reaction mixture was added to each well of the RapidScan PCR plate. The plate was placed in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94°C for 3min.) followed by 35 cycles of [(94°C for 45 sec.) (53°C for 2 min.) (72°C for 45 sec.)]. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel stained with ethidium bromide.

nGPCR-56 was expressed in peripheral blood lymphocytes, testis, salivary gland, kidney, spleen, skin, stomach, placenta, ovary, bone marrow, fetal liver, small intestine, and fetal brain.

Expression of nGPCR-56 in the brain provides an indication that modulators of nGPCR-56 activity have utility for treating neurological disorders, including but not limited to, movement disorders, affective disorders, metabolic disorders, inflammatory disorders and cancers. Use of nGPCR-56 modulators, including

nGPCR-56 ligands and anti-nGPCR-56 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

nGPCR-58

The RapidScan™ Gene Expression Panel was used to generate a comprehensive expression profile of the putative GPCR in human tissues. Human tissues in the array included: brain, heart, kidney, spleen, liver, lung, small intestine, muscle, testis, ovary, prostate, and PBL. Human brain regions in the array included: frontal lobe, temporal lobe, cerebellum, hippocampus, substantia nigra, caudate nucleus, amygdala, thalamus, hypothalamus, pons, medulla and spinal cord.

Expression of the nGPCR-58 in the various tissues was detected by using PCR primers designed based on the available sequence of the receptor that will prime the synthesis of a 282bp fragment in the presence of the appropriate cDNA. The forward primer was:

CAGAGCTTGATGATGAGGAC (SEQ ID NO: 127), and the reverse primer was:

CCCATAGGAAGTAGTAGAAG (SEQ ID NO: 128).

The PCR reaction mixture was added to each well of the PCR plate. The plate was placed in a GeneAmp PCR9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The plate was then exposed to the following cycling parameters: Pre-soak 94° for 3 min; denaturation at 94° for 30 seconds; annealing at primer T_m for 45 seconds; extension at 72° for 2 minutes; for 35 cycles. PCR productions were then separated and analyzed by electrophoresis on a 1.5-% agarose gel.

The 4-log dilution range of cDNA deposited on the plate ensured that the amplification reaction was within the linear range and, hence, facilitated semi-quantitative determination of relative mRNA accumulation in the various tissues or brain regions examined.

nGPCR-58 was expressed in all tissues included on the array, including brain, muscle, prostate, kidney, peripheral blood lymphocytes, liver, lung, small intestine, spleen, testis, heart, and ovary. Within the brain, nGPCR-58 was expressed in many regions including, but not limited to cerebellum, substantia nigra, thalamus, pons, spinal cord, frontal lobe, temporal lobe, hippocampus, caudate nucleus, amygdala, hypothalamus, and medulla.

Expression of the nGPCR-58 in the brain provided an indication that modulators of nGPCR-58 activity have utility for treating disorders, including but not

limited to, schizophrenia, affective disorders, ADHD/ADD (*i.e.*, Attention Deficit-Hyperactivity Disorder/Attention Deficit Disorder), neural disorders such as Alzheimer's disease, Parkinson's disease, migraine, senile dementia, depression, anxiety, bipolar disease, epilepsy, neuritis, neurasthenia, neuropathy, neuroses, metabolic disorders, inflammatory disorders, cancers and the like. Use of nGPCR-58 modulators, including nGPCR-58 ligands and anti-nGPCR-58 antibodies, to treat individuals having such disease states is intended as an aspect of the invention.

EXAMPLE 5: NORTHERN BLOT ANALYSIS

Northern blots are performed to examine the expression of nGPCR-x mRNA. The sense orientation oligonucleotide and the antisense-orientation oligonucleotide, described above, are used as primers to amplify a portion of the GPCR-x cDNA sequence of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185.

Multiple human tissue northern blots from Clontech (Human II # 7767-1) are hybridized with the probe. Pre-hybridization is carried out at 42 C for 4 hours in 5xSSC, 1X Denhardt's reagent, 0.1% SDS, 50% formamide, 250 mg/ml salmon sperm DNA. Hybridization is performed overnight at 42°C in the same mixture with the addition of about 1.5×10^6 cpm/ml of labeled probe.

The probe is labeled with α-^{32}P-dCTP by Rediprime™ DNA labeling system (Amersham Pharmacia), purified on Nick Column™ (Amersham Pharmacia) and added to the hybridization solution. The filters are washed several times at 42 C in 0.2x SSC, 0.1% SDS. Filters are exposed to Kodak XAR film (Eastman Kodak Company, Rochester, N.Y., USA) with intensifying screen at –80°C.

EXAMPLE 6: RECOMBINANT EXPRESSION OF nGPCR-X IN EUKARYOTIC HOST CELLS

A. Expression of nGPCR-x in Mammalian Cells

To produce nGPCR-x protein, a nGPCR-x-encoding polynucleotide is expressed in a suitable host cell using a suitable expression vector and standard genetic engineering techniques. For example, the nGPCR-x-encoding sequence described in Example 1 is subcloned into the commercial expression vector pzeoSV2 (Invitrogen, San Diego, CA) and transfected into Chinese Hamster Ovary (CHO) cells

using the transfection reagent FuGENE6™ (Boehringer-Mannheim) and the transfection protocol provided in the product insert. Other eukaryotic cell lines, including human embryonic kidney (HEK 293) and COS cells, are suitable as well. Cells stably expressing nGPCR-x are selected by growth in the presence of 100 μg/ml zeocin (Stratagene, LaJolla, CA). Optionally, nGPCR-x may be purified from the cells using standard chromatographic techniques. To facilitate purification, antisera is raised against one or more synthetic peptide sequences that correspond to portions of the nGPCR-x amino acid sequence, and the antisera is used to affinity purify nGPCR-x. The nGPCR-x also may be expressed in-frame with a tag sequence (*e.g.*, polyhistidine, hemagluttinin, FLAG) to facilitate purification. Moreover, it will be appreciated that many of the uses for nGPCR-x polypeptides, such as assays described below, do not require purification of nGPCR-x from the host cell.

B. Expression of nGPCR-x in 293 cells

For expression of nGPCR-x in mammalian cells 293 (transformed human, primary embryonic kidney cells), a plasmid bearing the relevant nGPCR-x coding sequence is prepared, using vector pSecTag2A (Invitrogen). Vector pSecTag2A contains the murine IgK chain leader sequence for secretion, the c-myc epitope for detection of the recombinant protein with the anti-myc antibody, a C-terminal polyhistidine for purification with nickel chelate chromatography, and a Zeocin resistant gene for selection of stable transfectants. The forward primer for amplification of this GPCR cDNA is determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce the *HindIII* cloning site and nucleotides matching the GPCR sequence. The reverse primer is also determined by routine procedures and preferably contains a 5' extension of nucleotides to introduce an *XhoI* restriction site for cloning and nucleotides corresponding to the reverse complement of the nGPCR-x sequence. The PCR conditions are 55°C as the annealing temperature. The PCR product is gel purified and cloned into the *HindIII-XhoI* sites of the vector.

The DNA is purified using Qiagen chromatography columns and transfected into 293 cells using DOTAP™ transfection media (Boehringer Mannheim, Indianapolis, IN). Transiently transfected cells are tested for expression after 24 hours of transfection, using western blots probed with anti-His and anti-nGPCR-x

peptide antibodies. Permanently transfected cells are selected with Zeocin and propagated. Production of the recombinant protein is detected from both cells and media by western blots probed with anti-His, anti-Myc or anti-GPCR peptide antibodies.

C. Expression of nGPCR-x in COS cells

For expression of the nGPCR-x in COS7 cells, a polynucleotide molecule having an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 can be cloned into vector p3-CI. This vector is a pUC18-derived plasmid that contains the HCMV (human cytomegalovirus) promoter-intron located upstream from the bGH (bovine growth hormone) polyadenylation sequence and a multiple cloning site. In addition, the plasmid contains the dhrf (dihydrofolate reductase) gene which provides selection in the presence of the drug methotrexane (MTX) for selection of stable transformants.

The forward primer is determined by routine procedures and preferably contains a 5' extension which introduces an *XbaI* restriction site for cloning, followed by nucleotides which correspond to an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. The reverse primer is also determined by routine procedures and preferably contains 5'- extension of nucleotides which introduces a *SalI* cloning site followed by nucleotides which correspond to the reverse complement of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. The PCR consists of an initial denaturation step of 5 min at 95°C 30 cycles of 30 sec denaturation at 95°C, 30 sec annealing at 58°C and 30 sec extension at 72°C, followed by 5 min extension at 72°C. The PCR product is gel purified and ligated into the *XbaI* and *SalI* sites of vector p3-CI. This construct is transformed into *E. coli* cells for amplification and DNA purification. The DNA is purified with Qiagen chromatography columns and transfected into COS 7 cells using Lipofectamine™ reagent from BRL, following the manufacturer's protocols. Forty-eight and 72 hours after transfection, the media and the cells are tested for recombinant protein expression.

nGPCR-x expressed from a COS cell culture can be purified by concentrating the cell-growth media to about 10 mg of protein/ml, and purifying the protein by, for

example, chromatography. Purified nGPCR-x is concentrated to 0.5 mg/ml in an Amicon concentrator fitted with a YM-10 membrane and stored at -80°C.

D. Expression of nGPCR-x in Insect Cells

For expression of nGPCR-x in a baculovirus system, a polynucleotide molecule having an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 can be amplified by PCR. The forward primer is determined by routine procedures and preferably contains a 5' extension which adds the *NdeI* cloning site, followed by nucleotides which correspond to an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185. The reverse primer is also determined by routine procedures and preferably contains a 5' extension which introduces the *KpnI* cloning site, followed by nucleotides which correspond to the reverse complement of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185.

The PCR product is gel purified, digested with *NdeI* and *KpnI*, and cloned into the corresponding sites of vector pACHTL-A (Pharmingen, San Diego, CA). The pAcHTL expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV), and a 6XHis tag upstream from the multiple cloning site. A protein kinase site for phosphorylation and a thrombin site for excision of the recombinant protein precede the multiple cloning site is also present. Of course, many other baculovirus vectors could be used in place of pAcHTL-A, such as pAc373, pVL941 and pAcIM1. Other suitable vectors for the expression of GPCR polypeptides can be used, provided that the vector construct includes appropriately located signals for transcription, translation, and trafficking, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow *et al.*, Virology 170:31-39, among others.

The virus is grown and isolated using standard baculovirus expression methods, such as those described in Summers *et al.* (A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experimental Station Bulletin No. 1555 (1987)).

In a preferred embodiment, pAcHLT-A containing nGPCR-x gene is introduced into baculovirus using the "BaculoGold™" transfection kit (Pharmingen,

San Diego, CA) using methods established by the manufacturer. Individual virus isolates are analyzed for protein production by radiolabeling infected cells with ^{35}S-methionine at 24 hours post infection. Infected cells are harvested at 48 hours post infection, and the labeled proteins are visualized by SDS-PAGE. Viruses exhibiting high expression levels can be isolated and used for scaled up expression.

For expression of a nGPCR-x polypeptide in a Sf9 cells, a polynucleotide molecule having the sequence of an odd numbered nucleotide sequence ranging from SEQ ID NO: 1 to SEQ ID NO: 93 and SEQ ID NO: 185 can be amplified by PCR using the primers and methods described above for baculovirus expression. The nGPCR-x cDNA is cloned into vector pAcHLT-A (Pharmingen) for expression in Sf9 insect. The insert is cloned into the *NdeI* and *KpnI* sites, after elimination of an internal *NdeI* site (using the same primers described above for expression in baculovirus). DNA is purified with Qiagen chromatography columns and expressed in Sf9 cells. Preliminary Western blot experiments from non-purified plaques are tested for the presence of the recombinant protein of the expected size which reacted with the GPCR-specific antibody. These results are confirmed after further purification and expression optimization in HiG5 cells.

EXAMPLE 7: INTERACTION TRAP/TWO-HYBRID SYSTEM

In order to assay for nGPCR-x-interacting proteins, the interaction trap/two-hybrid library screening method can be used. This assay was first described in Fields *et al.*, *Nature,* **1989**, *340*, 245, which is incorporated herein by reference in its entirety. A protocol is published in Current Protocols in Molecular Biology 1999, John Wiley & Sons, NY, and Ausubel, F. M. *et al.* 1992, Short protocols in molecular biology, Fourth edition, Greene and Wiley-interscience, NY, each of which is incorporated herein by reference in its entirety. Kits are available from Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

A fusion of the nucleotide sequences encoding all or partial nGPCR-x and the yeast transcription factor GAL4 DNA-binding domain (DNA-BD) is constructed in an appropriate plasmid (*i.e.*, pGBKT7) using standard subcloning techniques. Similarly, a GAL4 active domain (AD) fusion library is constructed in a second plasmid (*i.e.*, pGADT7) from cDNA of potential GPCR-binding proteins (for protocols on forming cDNA libraries, see Sambrook *et al.* 1989, Molecular cloning: a laboratory manual,

second edition, Cold Spring Harbor Press, Cold Spring Harbor, NY), which is incorporated herein by reference in its entirety. The DNA-BD/nGPCR-x fusion construct is verified by sequencing, and tested for autonomous reporter gene activation and cell toxicity, both of which would prevent a successful two-hybrid analysis. Similar controls are performed with the AD/library fusion construct to ensure expression in host cells and lack of transcriptional activity. Yeast cells are transformed (*ca.* 105 transformants/mg DNA) with both the nGPCR-x and library fusion plasmids according to standard procedures (Ausubel *et al.*, 1992, Short protocols in molecular biology, fourth edition, Greene and Wiley-interscience, NY, which is incorporated herein by reference in its entirety). *In vivo* binding of DNA-BD/nGPCR-x with AD/library proteins results in transcription of specific yeast plasmid reporter genes (*i.e.*, lacZ, HIS3, ADE2, LEU2). Yeast cells are plated on nutrient-deficient media to screen for expression of reporter genes. Colonies are dually assayed for β-galactosidase activity upon growth in Xgal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) supplemented media (filter assay for β-galactosidase activity is described in Breeden *et al.*, Cold Spring Harb. Symp. Quant. Biol., 1985, 50, 643, which is incorporated herein by reference in its entirety). Positive AD-library plasmids are rescued from transformants and reintroduced into the original yeast strain as well as other strains containing unrelated DNA-BD fusion proteins to confirm specific nGPCR-x/library protein interactions. Insert DNA is sequenced to verify the presence of an open reading frame fused to GAL4 AD and to determine the identity of the nGPCR-x-binding protein.

EXAMPLE 8: MOBILITY SHIFT DNA-BINDING ASSAY USING GEL ELECTROPHORESIS

A gel electrophoresis mobility shift assay can rapidly detect specific protein-DNA interactions. Protocols are widely available in such manuals as Sambrook *et al.* **1989**, *Molecular cloning: a laboratory* manual, second edition, Cold Spring Harbor Press, Cold Spring Harbor, NY and Ausubel, F. M. *et al.*, 1992, Short Protocols in Molecular Biology, fourth edition, Greene and Wiley-interscience, NY, each of which is incorporated herein by reference in its entirety.

Probe DNA(<300 bp) is obtained from synthetic oligonucleotides, restriction endonuclease fragments, or PCR fragments and end-labeled with ^{32}P. An aliquot of

purified nGPCR-x (*ca.* 15 μg) or crude nGPCR-x extract (*ca.* 15 ng) is incubated at constant temperature (in the range 22-37 C) for at least 30 minutes in 10-15 μl of buffer (*i.e.* TAE or TBE, pH 8.0-8.5) containing radiolabeled probe DNA, nonspecific carrier DNA (*ca.* 1 μg), BSA (300 μg/ml), and 10% (v/v) glycerol. The reaction mixture is then loaded onto a polyacrylamide gel and run at 30-35 mA until good separation of free probe DNA from protein-DNA complexes occurs. The gel is then dried and bands corresponding to free DNA and protein-DNA complexes are detected by autoradiography.

EXAMPLE 9: ANTIBODIES TO nGPCR-X

Standard techniques are employed to generate polyclonal or monoclonal antibodies to the nGPCR-x receptor, and to generate useful antigen-binding fragments thereof or variants thereof, including "humanized" variants. Such protocols can be found, for example, in Sambrook *et al.* (1989) and Harlow *et al.* (Eds.), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY (1988). In one embodiment, recombinant nGPCR-x polypeptides (or cells or cell membranes containing such polypeptides) are used as antigen to generate the antibodies. In another embodiment, one or more peptides having amino acid sequences corresponding to an immunogenic portion of nGPCR-x (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more amino acids) are used as antigen. Peptides corresponding to extracellular portions of nGPCR-x, especially hydrophilic extracellular portions, are preferred. The antigen may be mixed with an adjuvant or linked to a hapten to increase antibody production.

A. Polyclonal or Monoclonal antibodies

As one exemplary protocol, recombinant nGPCR-x or a synthetic fragment thereof is used to immunize a mouse for generation of monoclonal antibodies (or larger mammal, such as a rabbit, for polyclonal antibodies). To increase antigenicity, peptides are conjugated to Keyhole Lympet Hemocyanin (Pierce), according to the manufacturer's recommendations. For an initial injection, the antigen is emulsified with Freund's Complete Adjuvant and injected subcutaneously. At intervals of two to three weeks, additional aliquots of nGPCR-x antigen are emulsified with Freund's Incomplete Adjuvant and injected subcutaneously. Prior to the final booster injection, a serum sample is taken from the immunized mice and assayed by western blot to

confirm the presence of antibodies that immunoreact with nGPCR-x. Serum from the immunized animals may be used as polyclonal antisera or used to isolate polyclonal antibodies that recognize nGPCR-x. Alternatively, the mice are sacrificed and their spleen removed for generation of monoclonal antibodies.

To generate monoclonal antibodies, the spleens are placed in 10 ml serum-free RPMI 1640, and single cell suspensions are formed by grinding the spleens in serum-free RPMI 1640, supplemented with 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units/ml penicillin, and 100 μg/ml streptomycin (RPMI) (Gibco, Canada). The cell suspensions are filtered and washed by centrifugation and resuspended in serum-free RPMI. Thymocytes taken from three naive Balb/c mice are prepared in a similar manner and used as a Feeder Layer. NS-1 myeloma cells, kept in log phase in RPMI with 10% fetal bovine serum (FBS) (Hyclone Laboratories, Inc., Logan, Utah) for three days prior to fusion, are centrifuged and washed as well.

To produce hybridoma fusions, spleen cells from the immunized mice are combined with NS-1 cells and centrifuged, and the supernatant is aspirated. The cell pellet is dislodged by tapping the tube, and 2 ml of 37°C PEG 1500 (50% in 75 mM HEPES, pH 8.0) (Boehringer-Mannheim) is stirred into the pellet, followed by the addition of serum-free RPMI. Thereafter, the cells are centrifuged, resuspended in RPMI containing 15% FBS, 100 μM sodium hypoxanthine, 0.4 μM aminopterin, 16 μM thymidine (HAT) (Gibco), 25 units/ml IL-6 (Boehringer-Mannheim) and 1.5 x 10^6 thymocytes/ml, and plated into 10 Corning flat-bottom 96-well tissue culture plates (Corning, Corning New York).

On days 2, 4, and 6 after the fusion, 100 μl of medium is removed from the wells of the fusion plates and replaced with fresh medium. On day 8, the fusions are screened by ELISA, testing for the presence of mouse IgG that binds to nGPCR-x. Selected fusion wells are further cloned by dilution until monoclonal cultures producing anti-nGPCR-x antibodies are obtained.

B. Humanization of anti-nGPCR-x monoclonal antibodies

The expression pattern of nGPCR-x as reported herein and the proven track record of GPCRs as targets for therapeutic intervention suggest therapeutic indications for nGPCR-x inhibitors (antagonists). nGPCR-x-neutralizing antibodies comprise one class of therapeutics useful as nGPCR-x antagonists. Following are

protocols to improve the utility of anti-nGPCR-x monoclonal antibodies as therapeutics in humans by "humanizing" the monoclonal antibodies to improve their serum half-life and render them less immunogenic in human hosts (*i.e.*, to prevent human antibody response to non-human anti-nGPCR-x antibodies).

The principles of humanization have been described in the literature and are facilitated by the modular arrangement of antibody proteins. To minimize the possibility of binding complement, a humanized antibody of the IgG4 isotype is preferred.

For example, a level of humanization is achieved by generating chimeric antibodies comprising the variable domains of non-human antibody proteins of interest with the constant domains of human antibody molecules. (See, *e.g.*, Morrison *et al.*, Adv. Immunol., 44:65-92 (1989)). The variable domains of nGPCR-x-neutralizing anti-nGPCR-x antibodies are cloned from the genomic DNA of a B-cell hybridoma or from cDNA generated from mRNA isolated from the hybridoma of interest. The V region gene fragments are linked to exons encoding human antibody constant domains, and the resultant construct is expressed in suitable mammalian host cells (*e.g.*, myeloma or CHO cells).

To achieve an even greater level of humanization, only those portions of the variable region gene fragments that encode antigen-binding complementarity determining regions ("CDR") of the non-human monoclonal antibody genes are cloned into human antibody sequences. (See, *e.g.*, Jones *et al.*, Nature 321:522-525 (1986); Riechmann *et al.*, Nature 332:323-327 (1988); Verhoeyen *et al.*, Science 239:1534-36 (1988); and Tempest *et al.*, Bio/Technology 9: 266-71 (1991)). If necessary, the β-sheet framework of the human antibody surrounding the CDR3 regions also is modified to more closely mirror the three dimensional structure of the antigen-binding domain of the original monoclonal antibody. (See Kettleborough *et al.*, Protein Engin., 4:773-783 (1991); and Foote *et al.*, J. Mol. Biol., 224:487-499 (1992)).

In an alternative approach, the surface of a non-human monoclonal antibody of interest is humanized by altering selected surface residues of the non-human antibody, *e.g.*, by site-directed mutagenesis, while retaining all of the interior and contacting residues of the non-human antibody. See Padlan, Molecular Immunol., 28(4/5):489-98 (1991).

The foregoing approaches are employed using nGPCR-x-neutralizing anti-nGPCR-x monoclonal antibodies and the hybridomas that produce them to generate humanized nGPCR-x-neutralizing antibodies useful as therapeutics to treat or palliate conditions wherein nGPCR-x expression or ligand-mediated nGPCR-x signaling is detrimental.

C. Human nGPCR-x-Neutralizing Antibodies from Phage Display

Human nGPCR-x-neutralizing antibodies are generated by phage display techniques such as those described in Aujame *et al.*, Human Antibodies 8(4):155-168 (1997); Hoogenboom, TIBTECH 15:62-70 (1997); and Rader *et al.*, Curr. Opin. Biotechnol. 8:503-508 (1997), all of which are incorporated by reference. For example, antibody variable regions in the form of Fab fragments or linked single chain Fv fragments are fused to the amino terminus of filamentous phage minor coat protein pIII. Expression of the fusion protein and incorporation thereof into the mature phage coat results in phage particles that present an antibody on their surface and contain the genetic material encoding the antibody. A phage library comprising such constructs is expressed in bacteria, and the library is screened for nGPCR-x-specific phage-antibodies using labeled or immobilized nGPCR-x as antigen-probe.

D. Human nGPCR-x-neutralizing antibodies from transgenic mice

Human nGPCR-x-neutralizing antibodies are generated in transgenic mice essentially as described in Bruggemann *et al.*, Immunol. Today 17(8):391-97 (1996) and Bruggemann *et al.*, Curr. Opin. Biotechnol. 8:455-58 (1997). Transgenic mice carrying human V-gene segments in germline configuration and that express these transgenes in their lymphoid tissue are immunized with a nGPCR-x composition using conventional immunization protocols. Hybridomas are generated using B cells from the immunized mice using conventional protocols and screened to identify hybridomas secreting anti-nGPCR-x human antibodies (*e.g.*, as described above).

EXAMPLE 10: ASSAYS TO IDENTIFY MODULATORS OF nGPCR-X ACTIVITY

Set forth below are several nonlimiting assays for identifying modulators (agonists and antagonists) of nGPCR-x activity. Among the modulators that can be

identified by these assays are natural ligand compounds of the receptor; synthetic analogs and derivatives of natural ligands; antibodies, antibody fragments, and/or antibody-like compounds derived from natural antibodies or from antibody-like combinatorial libraries; and/or synthetic compounds identified by high-throughput screening of libraries; and the like. All modulators that bind nGPCR-x are useful for identifying nGPCR-x in tissue samples (*e.g.*, for diagnostic purposes, pathological purposes, and the like). Agonist and antagonist modulators are useful for up-regulating and down-regulating nGPCR-x activity, respectively, to treat disease states characterized by abnormal levels of nGPCR-x activity. The assays may be performed using single putative modulators, and/or may be performed using a known agonist in combination with candidate antagonists (or visa versa).

A. cAMP Assays

In one type of assay, levels of cyclic adenosine monophosphate (cAMP) are measured in nGPCR-x-transfected cells that have been exposed to candidate modulator compounds. Protocols for cAMP assays have been described in the literature. (See, *e.g.*, Sutherland *et al*, Circulation 37: 279 (1968); Frandsen *et al.*, Life Sciences 18: 529-541 (1976); Dooley *et al.*, Journal of Pharmacology and Experimental Therapeutics 283 (2): 735-41 (1997); and George *et al.*, Journal of Biomolecular Screening 2 (4): 235-40 (1997)). An exemplary protocol for such an assay, using an Adenylyl Cyclase Activation FlashPlate® Assay from NEN™ Life Science Products, is set forth below.

Briefly, the nGPCR-x coding sequence (*e.g.*, a cDNA or intronless genomic DNA) is subcloned into a commercial expression vector, such as pzeoSV2 (Invitrogen), and transiently transfected into Chinese Hamster Ovary (CHO) cells using known methods, such as the transfection protocol provided by Boehringer-Mannheim when supplying the FuGENE 6 transfection reagent. Transfected CHO cells are seeded into 96-well microplates from the FlashPlate® assay kit, which are coated with solid scintillant to which antisera to cAMP has been bound. For a control, some wells are seeded with wild type (untransfected) CHO cells. Other wells in the plate receive various amounts of a cAMP standard solution for use in creating a standard curve.

One or more test compounds (*i.e.*, candidate modulators) are added to the cells in each well, with water and/or compound-free medium/diluent serving as a control or

controls. After treatment, cAMP is allowed to accumulate in the cells for exactly 15 minutes at room temperature. The assay is terminated by the addition of lysis buffer containing [^{125}I]-labeled cAMP, and the plate is counted using a Packard Topcount™ 96-well microplate scintillation counter. Unlabeled cAMP from the lysed cells (or from standards) and fixed amounts of [^{125}I]-cAMP compete for antibody bound to the plate. A standard curve is constructed, and cAMP values for the unknowns are obtained by interpolation. Changes in intracellular cAMP levels of cells in response to exposure to a test compound are indicative of nGPCR-x modulating activity. Modulators that act as agonists of receptors which couple to the G_s subtype of G proteins will stimulate production of cAMP, leading to a measurable 3-10 fold increase in cAMP levels. Agonists of receptors which couple to the $G_{i/o}$ subtype of G proteins will inhibit forskolin-stimulated cAMP production, leading to a measurable decrease in cAMP levels of 50-100%. Modulators that act as inverse agonists will reverse these effects at receptors that are either constitutively active or activated by known agonists.

B. Aequorin Assays

In another assay, cells (*e.g.*, CHO cells) are transiently co-transfected with both a nGPCR-x expression construct and a construct that encodes the photoprotein apoaquorin. In the presence of the cofactor coelenterazine, apoaquorin will emit a measurable luminescence that is proportional to the amount of intracellular (cytoplasmic) free calcium. (See generally, Cobbold, *et al.* "Aequorin measurements of cytoplasmic free calcium," *In*: McCormack J.G. and Cobbold P.H., eds., *Cellular Calcium: A Practical Approach*. Oxford:IRL Press (1991); Stables *et al.*, Analytical Biochemistry 252: 115-26 (1997); and Haugland, Handbook of Fluorescent Probes and Research Chemicals. Sixth edition. Eugene OR: Molecular Probes (1996).)

In one exemplary assay, nGPCR-x is subcloned into the commercial expression vector pzeoSV2 (Invitrogen) and transiently co-transfected along with a construct that encodes the photoprotein apoaquorin (Molecular Probes, Eugene, OR) into CHO cells using the transfection reagent FuGENE 6 (Boehringer-Mannheim) and the transfection protocol provided in the product insert.

The cells are cultured for 24 hours at 37°C in MEM (Gibco/BRL, Gaithersburg, MD) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 μg/ml streptomycin, at which time the medium is changed to

serum-free MEM containing 5 μM coelenterazine (Molecular Probes, Eugene, OR). Culturing is then continued for two additional hours at 37°C. Subsequently, cells are detached from the plate using VERSEN (Gibco/BRL), washed, and resuspended at 200,000 cells/ml in serum-free MEM.

Dilutions of candidate nGPCR-x modulator compounds are prepared in serum-free MEM and dispensed into wells of an opaque 96-well assay plate at 50 μl/well. Plates are then loaded onto an MLX microtiter plate luminometer (Dynex Technologies, Inc., Chantilly, VA). The instrument is programmed to dispense 50 μl cell suspensions into each well, one well at a time, and immediately read luminescence for 15 seconds. Dose-response curves for the candidate modulators are constructed using the area under the curve for each light signal peak. Data are analyzed with SlideWrite, using the equation for a one-site ligand, and EC_{50} values are obtained. Changes in luminescence caused by the compounds are considered indicative of modulatory activity. Modulators that act as agonists at receptors which couple to the G_q subtype of G proteins give an increase in luminescence of up to 100 fold. Modulators that act as inverse agonists will reverse this effect at receptors that are either constitutively active or activated by known agonists.

C. <u>Luciferase Reporter Gene Assay</u>

The photoprotein luciferase provides another useful tool for assaying for modulators of nGPCR-x activity. Cells (*e.g.*, CHO cells or COS 7 cells) are transiently co-transfected with both a nGPCR-x expression construct (*e.g.*, nGPCR-x in pzeoSV2) and a reporter construct which includes a gene for the luciferase protein downstream from a transcription factor binding site, such as the cAMP-response element (CRE), AP-1, or NF-kappa B. Agonist binding to receptors coupled to the G_s subtype of G proteins leads to increases in cAMP, thereby activating the CRE transcription factor and resulting in expression of the luciferase gene. Agonist binding to receptors coupled to the G_q subtype of G protein leads to production of diacylglycerol that activates protein kinase C, which activates the AP-1 or NF-kappa B transcription factors, in turn resulting in expression of the luciferase gene. Expression levels of luciferase reflect the activation status of the signaling events. (See generally, George *et al.*, Journal of Biomolecular Screening 2(4): 235-240 (1997); and Stratowa *et al.*, Current Opinion in Biotechnology 6: 574-581 (1995)).

Luciferase activity may be quantitatively measured using, *e.g.*, luciferase assay reagents that are commercially available from Promega (Madison, WI).

In one exemplary assay, CHO cells are plated in 24-well culture dishes at a density of 100,000 cells/well one day prior to transfection and cultured at 37°C in MEM (Gibco/BRL) supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 μg/ml streptomycin. Cells are transiently co-transfected with both a nGPCR-x expression construct and a reporter construct containing the luciferase gene. The reporter plasmids CRE-luciferase, AP-1-luciferase and NF-kappaB-luciferase may be purchased from Stratagene (LaJolla, CA). Transfections are performed using the FuGENE 6 transfection reagent (Boehringer-Mannheim) according to the supplier's instructions. Cells transfected with the reporter construct alone are used as a control. Twenty-four hours after transfection, cells are washed once with PBS pre-warmed to 37°C. Serum-free MEM is then added to the cells either alone (control) or with one or more candidate modulators and the cells are incubated at 37°C for five hours. Thereafter, cells are washed once with ice-cold PBS and lysed by the addition of 100 μl of lysis buffer per well from the luciferase assay kit supplied by Promega. After incubation for 15 minutes at room temperature, 15 μl of the lysate is mixed with 50 μl of substrate solution (Promega) in an opaque-white, 96-well plate, and the luminescence is read immediately on a Wallace model 1450 MicroBeta scintillation and luminescence counter (Wallace Instruments, Gaithersburg, MD).

Differences in luminescence in the presence versus the absence of a candidate modulator compound are indicative of modulatory activity. Receptors that are either constitutively active or activated by agonists typically give a 3 to 20-fold stimulation of luminescence compared to cells transfected with the reporter gene alone. Modulators that act as inverse agonists will reverse this effect.

D. Intracellular calcium measurement using FLIPR

Changes in intracellular calcium levels are another recognized indicator of G protein-coupled receptor activity, and such assays can be employed to screen for modulators of nGPCR-x activity. For example, CHO cells stably transfected with a nGPCR-x expression vector are plated at a density of 4 x 10^4 cells/well in Packard black-walled, 96-well plates specially designed to discriminate fluorescence signals emanating from the various wells on the plate. The cells are incubated for 60 minutes

at 37°C in modified Dulbecco's PBS (D-PBS) containing 36 mg/L pyruvate and 1 g/L glucose with the addition of 1% fetal bovine serum and one of four calcium indicator dyes (Fluo-3™ AM, Fluo-4™ AM, Calcium Green™-1 AM, or Oregon Green™ 488 BAPTA-1 AM), each at a concentration of 4 μM. Plates are washed once with modified D-PBS without 1% fetal bovine serum and incubated for 10 minutes at 37°C to remove residual dye from the cellular membrane. In addition, a series of washes with modified D-PBS without 1% fetal bovine serum is performed immediately prior to activation of the calcium response.

A calcium response is initiated by the addition of one or more candidate receptor agonist compounds, calcium ionophore A23187 (10 μM; positive control), or ATP (4 μM; positive control). Fluorescence is measured by Molecular Device's FLIPR with an argon laser (excitation at 488 nm). (See, *e.g.*, Kuntzweiler *et al.*, Drug Development Research, 44(1):14-20 (1998)). The F-stop for the detector camera was set at 2.5 and the length of exposure was 0.4 milliseconds. Basal fluorescence of cells was measured for 20 seconds prior to addition of candidate agonist, ATP, or A23187, and the basal fluorescence level was subtracted from the response signal. The calcium signal is measured for approximately 200 seconds, taking readings every two seconds. Calcium ionophore A23187 and ATP increase the calcium signal 200% above baseline levels. In general, activated GPCRs increase the calcium signal approximately 10-15% above baseline signal.

E. <u>Mitogenesis Assay</u>

In a mitogenesis assay, the ability of candidate modulators to induce or inhibit nGPCR-x-mediated cell division is determined. (*See, e.g.*, Lajiness *et al.*, Journal of Pharmacology and Experimental Therapeutics 267(3): 1573-1581 (1993)). For example, CHO cells stably expressing nGPCR-x are seeded into 96-well plates at a density of 5000 cells/well and grown at 37°C in MEM with 10% fetal calf serum for 48 hours, at which time the cells are rinsed twice with serum-free MEM. After rinsing, 80 μl of fresh MEM, or MEM containing a known mitogen, is added along with 20 μl MEM containing varying concentrations of one or more candidate modulators or test compounds diluted in serum-free medium. As controls, some wells on each plate receive serum-free medium alone, and some receive medium containing 10% fetal bovine serum. Untransfected cells or cells transfected with vector alone also may serve as controls.

After culture for 16-18 hours, 1 μCi of [^{3}H]-thymidine (2 Ci/mmol) is added to the wells and cells are incubated for an additional 2 hours at 37°C. The cells are trypsinized and collected on filter mats with a cell harvester (Tomtec); the filters are then counted in a Betaplate counter. The incorporation of [^{3}H]-thymidine in serum-free test wells is compared to the results achieved in cells stimulated with serum (positive control). Use of multiple concentrations of test compounds permits creation and analysis of dose-response curves using the non-linear, least squares fit equation: $A = B \times [C/ (D + C)] + G$ where A is the percent of serum stimulation; B is the maximal effect minus baseline; C is the EC_{50}; D is the concentration of the compound; and G is the maximal effect. Parameters B, C and G are determined by Simplex optimization.

Agonists that bind to the receptor are expected to increase [^{3}H]-thymidine incorporation into cells, showing up to 80% of the response to serum. Antagonists that bind to the receptor will inhibit the stimulation seen with a known agonist by up to 100%.

F. [^{35}S]GTPγS Binding Assay

Because G protein-coupled receptors signal through intracellular G proteins whose activity involves GTP binding and hydrolysis to yield bound GDP, measurement of binding of the non-hydrolyzable GTP analog [^{35}S]GTPγS in the presence and absence of candidate modulators provides another assay for modulator activity. (*See, e.g.*, Kowal *et al.*, Neuropharmacology 37:179-187 (1998).)

In one exemplary assay, cells stably transfected with a nGPCR-x expression vector are grown in 10 cm tissue culture dishes to subconfluence, rinsed once with 5 ml of ice-cold Ca^{2+}/Mg^{2+}-free phosphate-buffered saline, and scraped into 5 ml of the same buffer. Cells are pelleted by centrifugation (500 x *g*, 5 minutes), resuspended in TEE buffer (25 mM Tris, pH 7.5 , 5 mM EDTA, 5 mM EGTA), and frozen in liquid nitrogen. After thawing, the cells are homogenized using a Dounce homogenizer (one ml TEE per plate of cells), and centrifuged at 1,000 x *g* for 5 minutes to remove nuclei and unbroken cells.

The homogenate supernatant is centrifuged at 20,000 x *g* for 20 minutes to isolate the membrane fraction, and the membrane pellet is washed once with TEE and resuspended in binding buffer (20 mM HEPES, pH 7.5, 150 mM NaCl, 10 mM

$MgCl_2$, 1 mM EDTA). The resuspended membranes can be frozen in liquid nitrogen and stored at -70°C until use.

Aliquots of cell membranes prepared as described above and stored at -70°C are thawed, homogenized, and diluted into buffer containing 20 mM HEPES, 10 mM $MgCl_2$, 1 mM EDTA, 120 mM NaCl, 10 μM GDP, and 0.2 mM ascorbate, at a concentration of 10-50 μg/ml. In a final volume of 90 μl, homogenates are incubated with varying concentrations of candidate modulator compounds or 100 μM GTP for 30 minutes at 30°C and then placed on ice. To each sample, 10 μl guanosine 5'-O-(3[^{35}S]thio) triphosphate (NEN, 1200 Ci/mmol; [^{35}S]-GTPγS), was added to a final concentration of 100-200 pM. Samples are incubated at 30°C for an additional 30 minutes, 1 ml of 10 mM HEPES, pH 7.4, 10 mM $MgCl_2$, at 4°C is added and the reaction is stopped by filtration.

Samples are filtered over Whatman GF/B filters and the filters are washed with 20 ml ice-cold 10 mM HEPES, pH 7.4, 10 mM $MgCl_2$. Filters are counted by liquid scintillation spectroscopy. Nonspecific binding of [^{35}S]-GTPγS is measured in the presence of 100 μM GTP and subtracted from the total. Compounds are selected that modulate the amount of [^{35}S]-GTPγS binding in the cells, compared to untransfected control cells. Activation of receptors by agonists gives up to a five-fold increase in [^{35}S]GTPγS binding. This response is blocked by antagonists.

G. <u>MAP Kinase Activity Assay</u>

Evaluation of MAP kinase activity in cells expressing a GPCR provides another assay to identify modulators of GPCR activity. (See, *e.g.*, Lajiness *et al.*, Journal of Pharmacology and Experimental Therapeutics 267(3):1573-1581 (1993) and Boulton *et al.*, Cell 65:663-675 (1991).)

In one embodiment, CHO cells stably transfected with nGPCR-x are seeded into 6-well plates at a density of 70,000 cells/well 48 hours prior to the assay. During this 48-hour period, the cells are cultured at 37°C in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 μg/ml streptomycin. The cells are serum-starved for 1-2 hours prior to the addition of stimulants.

For the assay, the cells are treated with medium alone or medium containing either a candidate agonist or 200 nM Phorbol ester- myristoyl acetate (*i.e.*, PMA, a positive control), and the cells are incubated at 37°C for varying times. To stop the

reaction, the plates are placed on ice, the medium is aspirated, and the cells are rinsed with 1 ml of ice-cold PBS containing 1 mM EDTA. Thereafter, 200 μl of cell lysis buffer (12.5 mM MOPS, pH 7.3, 12.5 mM glycerophosphate, 7.5 mM $MgCl_2$, 0.5 mM EGTA, 0.5 mM sodium vanadate, 1 mM benzamidine, 1 mM dithiothreitol, 10 μg/ml leupeptin, 10 μg/ml aprotinin, 2 μg/ml pepstatin A, and 1 μM okadaic acid) is added to the cells. The cells are scraped from the plates and homogenized by 10 passages through a 23 3/4 G needle, and the cytosol fraction is prepared by centrifugation at 20,000 x *g* for 15 minutes.

Aliquots (5-10 μl containing 1-5 μg protein) of cytosol are mixed with 1 mM MAPK Substrate Peptide (APRTPGGRR (SEQ ID NO: 129), Upstate Biotechnology, Inc., N.Y.) and 50 μM [γ-^{32}P]ATP (NEN, 3000 Ci/mmol), diluted to a final specific activity of ~2000 cpm/pmol, in a total volume of 25 μl. The samples are incubated for 5 minutes at 30°C, and reactions are stopped by spotting 20 μl on 2 cm^2 squares of Whatman P81 phosphocellulose paper. The filter squares are washed in 4 changes of 1% H_3PO_4, and the squares are subjected to liquid scintillation spectroscopy to quantitate bound label. Equivalent cytosolic extracts are incubated without MAPK substrate peptide, and the bound label from these samples are subtracted from the matched samples with the substrate peptide. The cytosolic extract from each well is used as a separate point. Protein concentrations are determined by a dye binding protein assay (Bio-Rad Laboratories). Agonist activation of the receptor is expected to result in up to a five-fold increase in MAPK enzyme activity. This increase is blocked by antagonists.

H. <u>[^{3}H]Arachidonic Acid Release</u>

The activation of GPCRs also has been observed to potentiate arachidonic acid release in cells, providing yet another useful assay for modulators of GPCR activity. (*See, e.g.*, Kanterman *et al.*, Molecular Pharmacology 39:364-369 (1991).) For example, CHO cells that are stably transfected with a nGPCR-x expression vector are plated in 24-well plates at a density of 15,000 cells/well and grown in MEM medium supplemented with 10% fetal bovine serum, 2 mM glutamine, 10 U/ml penicillin and 10 μg/ml streptomycin for 48 hours at 37°C before use. Cells of each well are labeled by incubation with [^{3}H]-arachidonic acid (Amersham Corp., 210 Ci/mmol) at 0.5 μCi/ml in 1 ml MEM supplemented with 10 mM HEPES, pH 7.5, and 0.5% fatty-

acid-free bovine serum albumin for 2 hours at 37°C. The cells are then washed twice with 1 ml of the same buffer.

Candidate modulator compounds are added in 1 ml of the same buffer, either alone or with 10 μM ATP and the cells are incubated at 37°C for 30 minutes. Buffer alone and mock-transfected cells are used as controls. Samples (0.5 ml) from each well are counted by liquid scintillation spectroscopy. Agonists which activate the receptor will lead to potentiation of the ATP-stimulated release of [^{3}H]-arachidonic acid. This potentiation is blocked by antagonists.

I. Extracellular Acidification Rate

In yet another assay, the effects of candidate modulators of nGPCR-x activity are assayed by monitoring extracellular changes in pH induced by the test compounds. (*See, e.g.*, Dunlop *et al.*, Journal of Pharmacological and Toxicological Methods 40(1):47-55 (1998).) In one embodiment, CHO cells transfected with a nGPCR-x expression vector are seeded into 12 mm capsule cups (Molecular Devices Corp.) at 4 x 10^5 cells/cup in MEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 10 U/ml penicillin, and 10 μg/ml streptomycin. The cells are incubated in this medium at 37°C in 5% CO_2 for 24 hours.

Extracellular acidification rates are measured using a Cytosensor microphysiometer (Molecular Devices Corp.). The capsule cups are loaded into the sensor chambers of the microphysiometer and the chambers are perfused with running buffer (bicarbonate-free MEM supplemented with 4 mM L-glutamine, 10 units/ml penicillin, 10 μg/ml streptomycin, 26 mM NaCl) at a flow rate of 100 μl/minute. Candidate agonists or other agents are diluted into the running buffer and perfused through a second fluid path. During each 60-second pump cycle, the pump is run for 38 seconds and is off for the remaining 22 seconds. The pH of the running buffer in the sensor chamber is recorded during the cycle from 43-58 seconds, and the pump is re-started at 60 seconds to start the next cycle. The rate of acidification of the running buffer during the recording time is calculated by the Cytosoft program. Changes in the rate of acidification are calculated by subtracting the baseline value (the average of 4 rate measurements immediately before addition of a modulator candidate) from the highest rate measurement obtained after addition of a modulator candidate. The selected instrument detects 61 mV/pH unit. Modulators that act as agonists of the receptor result in an increase in the rate of extracellular acidification compared to the

rate in the absence of agonist. This response is blocked by modulators which act as antagonists of the receptor.

EXAMPLE 11: *IN SITU* HYBRIDIZATION

DNA Probe Preparation For nGPCR-11, -16, -40, -54, and -56

DNA probes for *in situ* hybridization were prepared as follows. Two sets of primer pairs were prepared. The first set has the sequence for T7 polymerase promoter on the 5' primer to make the sense RNA, and the second set has the T7 polymerase promoter sequence on the 3' primer to make the antisense RNA. PCR was performed in a 50 µl reaction containing 36.5 µl H_2O, 5µl 10xTT buffer (140 mM Ammonium Sulfate, 0.1 % gelatine, 0.6 M Tris-tricine pH 8.4), 5 µl 25mM $MgCl_2$, 2 µl 10 mM dNTP, 0.4 µl Incyte clone 1722192 DNA, 0.5 µl AmpliTaq (PE Applied Biosystems), and 0.3 µl oligo1 (1 mg/ml) and 0.3 µl oligo2 (1mg/ml)[to make the sense RNA], or 0.3 µl oligo3 (1 mg/ml) and 0.3 µl oligo4 (1mg/ml)[to make the antisense RNA]. The PCR reaction involved one cycle at 94°C for 2 min followed by 35 cycles at 94°C for 30 sec, 60°C for 30 sec, 72°C for 30 sec. The two PCR reactions were loaded onto a 1.2 % agarose gel. The DNA band was excised from the gel, placed in a GenElute Agarose spin column (Supelco) and spun for 10 min at maximum speed. The eluted DNA was EtOH precipitated and resuspended in transcription buffer. The primer sequences for each nGPCR tested are listed below.

For nGPCR-11, the sense primers were:
GCGTAATACGACTCACTATAGGGAGACCGCGTGTCTGCTAGACTCTATTTC
C 3'(LW1658) (SEQ ID NO: 159), and:
5' TGCCACACTGATGCAACTCC 3' (LW1661) (SEQ ID NO: 160). The antisense primers were:
GCGTAATACGACTCACTATAGGGAGACCTGCCACACTGATGCAACTCC
(LW1659) SEQ ID NO: 161) and.
5'GCGTGTCTGCTAGACTCTATTTCC 3' (LW1660) (SEQ ID NO: 162). The primer pairs yielded a product of 275bp.

For nGPCR-16, the sense primers were:
5'GCGTAATACGACTCACTATAGGGAGACCGCACGCCACTCTTTACTATCC
C (LW1645) (SEQ ID NO: 163), and:

5' GCACAAAACACAATTCCATAAGCC 3' (LW1648) (SEQ ID NO: 164). The antisense primers were:
5'GCGTAATACGACTCACTATAGGGAGACCGCACAAAACACAATTCCATAAGCC 3' (LW1646) (SEQ ID NO: 165), and:
5' GCTACGCCACTCTTTACTATCCC 3'(LW1647) (SEQ ID NO: 166). The primer pairs yielded a product of 283 bp.

For nGPCR-40, the sense primers were:
5'GCGTAATACGACTCACTATAGGGAGACCTTATGAGCAGCAATTCATCCC 3'(LW1704) (SEQ ID NO: 167), and:
5'CACACCCACCAAGAAATCAG 3'(LW1707)(SEQ ID NO: 168). The antisense primers were:
5'GCGTAATACGACTCACTATAGGGAGACCCACACCCACCAAGAAATCAG 3'(LW1705) (SEQ ID NO: 169), and:
5' TTATGAGCAGCAATTCATCCC 3' (LW1706) (SEQ ID NO: 170). The primer pairs yielded a product of 251bp.

For nGPCR-54, the sense primers were:
5'GCGTAATACGACTCACTATAGGGAGACCCGATTATCCACACTTTGACCC 3' (LW1803) (SEQ ID NO: 171), and:
5' CTGAAAGTTGTCGCTGACC 3' (LW1634) (SEQ ID NO: 172). The anti-sense primers were:
GCGTAATACGACTCACTATAGGGAGACCCTGCTGAAAGTTGTCGCTGACC 3' (LW1804)(SEQ ID NO: 173), and:
5' CGATTATCCACACTTTGACCC 3' (LW1635) (SEQ ID NO: 174). The primer pairs yielded a product of 286 bp.

For nGPCR-56, the sense primers were:
GCGTAATACGACTCACTATAGGGAGACCCTGTAAAATTCACACAAGCACC 3' (LW1763) (SEQ ID NO: 175), and:
5'AGAAGACAGAGCAACCTCC 3' (LW1766) (SEQ ID NO: 176). The anti-sense primers were:
GCGTAATACGACTCACTATAGGGAGACCAGAAGACAGAGCAACCTCC (LW1764) (SEQ ID NO: 177) and:
CTGTAAAATTCACACAAGCACC (LW1765) (SEQ ID NO: 178). The primer pairs yielded a product of 272 bp.

DNA Probe Preparation For nGPCR-1

Probes for nGPCR-1 were prepared as above with the following modifications. Using a sense primer: GCATGGATCCTCTTTGCTGTATTTCACCCTC) (LW1595) (SEQ ID NO: 179) and an antisense primer: 5'GCATGAATTCACAATGCCAGTGATAAGGAAG 3' (LW1596) (SEQ ID NO: 180), a 271 bp fragment was generated by PCR. The fragment was digested with *BamHI* and *EcoRI* and ligated into a BluescriptII vector that had been cut with *BamHI* and *EcoRI*. The orientation of the insert was such that T7 polymerase generates the anti-sense strand and T3 polymerase generates the sense strand.

Histochemistry

Coronal and sagittal oriented rat brain sections were cryosectioned (20 µm thick) using a Reichert-Jung cryostat. The individual sections were thaw-mounted onto silanated, nuclease-free slides (CEL Associates, Inc., Houston, TX), and stored at -80°C. The sections were processed starting with post-fixation in cold 4% paraformaldehyde, rinsed in cold PBS, acetylated using acetic anhydride in triethanolamine buffer and dehydrated through 70%, 95%, and 100% alcohols at room temperature (RT). This was followed with delipidation in chloroform then rehydration in 100% and 95% alcohol at room temperature. Sections were air-dried prior to hybridization. Two PCR fragments (~ 250 bp) were generated, one that contained T7 polymerase on the 5' end (sense) and the other with T7 polymerase on the 3' end (antisense). The PCR fragments were labeled with ^{35}S-UTP to yield a specific activity of 0.655 x 10^6 cpm/pmol for antisense and 0.675 x 10^6 cpm/pmol for sense probe. Both riboprobes were denatured and added to hybridization buffer containing 50% formamide, 10% dextran, 0.3M NaCl, 10 mM Tris, 1 mM EDTA, 1X Denhardts, and 10 mM DTT. Sequential brain cryosections were hybridized with 45 µl/slide of the sense and antisense riboprobe hybridization mixture, then covered with silanized glass coverslips. The sections were hybridized overnight (15-18 hrs) at 42°C in an incubator.

Coverslips were washed off the slides in 1X SSC, followed by RNase A treatment, and high temperature stringency washes (3X, 20 mins at 41°C) in 0.1X SSC. Slides were dehydrated with 70%, 95% NH_4OAc, and 100% NH_4OAc alcohols, air-dried and exposed to Kodak BioMax MR-1 film. After 9 days of exposure, the

film was developed. This was followed with coating selected tissue slides with Kodak NTB-2 nuclear track emulsion and storing the slides in the dark for 23 days. The slides were then developed and counterstained with hematoxylin. Emulsion-coated sections were analyzed microscopically to determine the specificity of labeling. Presence of autoradiographic grains (generated by antisense probe hybridization) over cell bodies (versus between cell bodies) was used as an index of specific hybridization.

Results

In situ hybridization results indicated localization in the following brain areas:

nGPCR-1 was localized to the dentate gyrus of hippocampus, piriform cortex, and red nucleus.

nGPCR-11 was localized to the piriform cortex, hippocampus, red nucleus, subthalamic nuclei, dorsal raphe, interpeduncular nucleus, and habenula. nGPCR-16 was localized to the cortex, piriform cortex, hippocampus, thalamus, subthalamic nuclei, hypothalamus, bed nucleus stria terminalis and posterior striatum. nGPCR-40 was localized to the cortex, piriform cortex, hippocampus, substantia nigra compacta, hypothalamus, laterial septus, bed nucleus stria terminalis, thalamus, ventral tegmental area, interpeduncular nucleus, dorsal raphe, medical geniculate, islands of Calleja, subthamalmic nuclei, choroid plexus. nGPCR-54 was localized to the piriform cortex and hippocampus, including the dentate gyrus, CA1 and CA3. nGPCR-56 was localized to the piriform cortex, cortex, interpeduncular nuceus, red nucleus, hippocampus, habenula, substantia nigra pars compacta, mamillary body stria terminalis,hypothalamus, subthamalmic nuclei, corsal raphe, and ventral tegmental area.

EXAMPLE 12: CHROMOSOMAL LOCALIZATION

Methods

Chromosomal location of the genes encoding nGPCRs was determined using the Stanford G3 Radiation Hybrid Panel (Research Genetics, Inc., Huntsville, AL). This panel contains 83 radiation hybrid clones of the entire human genome created by the Stanford Human Genome Center. PCR reactions were assembled containing 25ng of DNA from each clone and the components of the Expand Hi-Fi PCR System™ (Roche Molecular Biochemicals, Indianapolis, IN) in a final reaction volume of 15 µl. PCR primers were synthesized by Genosys Corp., The Woodlands, TX. PCR

reactions were incubated in a GeneAmp 9700 PCR thermocycler (Perkin Elmer Applied Biosystems). The following cycling program was executed: Pre-soak at (94° for 3min.)(94° for 30 sec.)(52°C for 60 sec.)(72° for 2 min.)] for 35 cycles. PCR reaction products were then separated and analyzed by electrophoresis on a 2.0% agarose gel, and stained with ethidium bromide. Lanes were scored for the presence or absence of the expected PCR product and the results submitted to the Stanford Human Genome Center via e-mail for analysis (http://www-shgc.stanford.edu./RH/rhserverformnew.html).

nGPCR-40

PCR primers were designed based on the available sequence of the Celera sequence HUM_IDS|Contig|11000258115466. The forward primer used was:

5'ACAGCCCCAAAGCCAAACAC3' (SEQ ID NO: 181). The reverse primer was:

5'CCGCAGGAGCAATG-AAAATCAG3' (SEQ ID NO: 182). This primer set will prime the synthesis of a 220 base pair fragment in the presence of the appropriate genomic DNA.

G3 Radiation Hybrid Panel Analysis places nGPCR-40 on chromosome 6, most nearly linked to Stanford marker SHGC-1836 with a LOD score of 11.84. This marker lies at position 6q21. In a genome scanning data set, Cao *et al.* (Genomics 1997 Jul 1: 43(1): 1-8) found excess allele sharing for markers on 6q13-q26. Greatest allele sharing was at interval 6q21-q22.3 with a maximum multipoint MLS value of 3.06 close to marker D6S278. Replication data from a second data set found maximum multipoint MLS at the interval D6S424-D6S275. These results provide suggestive evidence for a susceptibility locus for schizophrenia in chromosome 6q from two independent data sets.

nGPCR-54

PCR primers were designed based on the available sequence of the Celera sequence GA_11824020. The forward primer used was:

5'CTGTCTCTCTGTCCTCTTCC3', (SEQ ID NO: 183). The reverse primer used was:

5'GCACCGATCTTCATTGAATTTC3', (SEQ ID NO: 184). This primer set will prime the synthesis of a 145 base pair fragment in the presence of the appropriate genomic DNA.

G3 Radiation Hybrid Panel Analysis places nGPCR-54 on chromosome 13, most nearly linked to Stanford marker SHGC-68276 with a LOD score of 6.31. This marker lies at position 13q32. Numerous investigations have found significant suggestion of linkage of schizophrenia to this region of chromosome 13q32. See, for example, Brzustowicz *et al.*, Am J Hum Genet 1999 Oct; 65(4): 1096-1103; Blouin *et al.*, Nat Genet 1998 Sep; 20(1): 70-3; Shaw *et al.*, Am J Med Genet. 1998 Sep 7; 81(5): 364-76; Lin *et al.*, Hum Genet 1997 Mar; 99(3): 417-20; Pulver *et al.*, Cold Spring Harb Symp Quant Biol 1996; 61:797-814.

Genes localized to chromosomal regions in linkage with schizophrenia are candidate genes for disease susceptibility. Genes in these regions with the potential to play a biochemical/functional role in the disease process (like G protein coupled receptors) have a high probability of being a disease-modifying locus. nGPCR-40 and -54, because of their chromosomal location, are attractive targets therefore for screening ligands useful in modulating cellular processes involved in schizophrenia.

EXAMPLE 13: CLONE DEPOSIT INFORMATION

In accordance with the Budapest Treaty, clones of the present invention have been deposited at the Agricultural Research Culture Collection (NRRL) International Depository Authority, 1815 N. University Street, Peoria, Illinois 61604, U.S.A. Accession numbers and deposit dates are provided below in Table 6.

Table 6: DEPOSIT INFORMATION

Clone	Accession Number NRRL	Budapest Treaty Deposit Date
nGPCR-1 (SEQ ID NO:73)	B-30243	2000 Jan 18
nGPCR -5 (SEQ ID NO: 75)	B-30244	2000 Jan 18
nGPCR -16 (SEQ ID NO: 81)	B-30245	2000 Jan 18
nGPCR -11 (SEQ ID NO: 79)	B-30258	2000 Feb 02
nGPCR -17 (SEQ ID NO: 23)	B-30259	2000 Feb 03
nGPCR -9 (SEQ ID NO: 77)	B-30262	2000 Feb 22
nGPCR -58 (SEQ ID NO: 91)	B-30274	2000 March 23
nGPCR -56 (SEQ ID NO: 89)	B-30288	2000 May 5
nGPCR -3 (SEQ ID NO:185)	B-30290	2000 May 5
nGPCR -54 (SEQ ID NO: 85)	B-30291	2000 May 5
nGPCR -40 (SEQ ID NO: 83*)	B-30299N	2000 June 02

* The clone deposited with NRLL Accession Number B30299N comprises a sequence identical to SEQ ID NO:83 but with the substitution of an "A" at nucleotide position 10.

Example 14 - Using nGPCR-x proteins to isolate neurotransmitters

The isolated nGPCR-x proteins, particularly nGPCR-1, nGPCR-3, nGPCR-9, nGPCR-11, nGPCR-16, nGPCR-40, nGPCR-54, nGPCR-56, and nGPCR-58, (SEQ ID NOS: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82, SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94, respectively) can be used to isolate novel or known neurotransmitters (Saito *et al.*, Nature 400: 265-269, 1999). The cDNAs that encode the isolated nGPCR-x can be cloned into mammalian expression vectors and used to stably or transiently transfect mammalian cells including CHO, Cos or HEK293 cells. Receptor expression can be determined by Northern blot analysis of transfected cells and identification of an appropriately sized mRNA band (predicted size from the cDNA). Brain regions shown by mRNA analysis to express each of the nGPCR-x proteins could be processed for peptide extraction using any of several protocols

((Reinsheidk R.K. *et al.*, Science 270: 243-247, 1996; Sakurai, T., *et al.*, Cell 92; 573-585, 1998; Hinuma, S., *et al.*, Nature 393: 272-276, 1998). Chromotographic fractions of brain extracts could be tested for ability to activate nGPCR-x proteins by measuring second messenger production such as changes in cAMP production in the presence or absence of forskolin, changes in inositol 3-phosphate levels, changes in intracellular calcium levels or by indirect measures of receptor activation including receptor stimulated mitogenesis, receptor mediated changes in extracellular acidification or receptor mediated changes in reporter gene activation in response to cAMP or calcium (these methods should all be referenced in other sections of the patent). Receptor activation could also be monitored by co-transfecting cells with a chimeric $GI_{q/i3}$ to force receptor coupling to a calcium stimulating pathway (Conklin *et al.*, Nature 363; 274-276, 1993). Neurotransmitter mediated activation of receptors could also be monitored by measuring changes in $[^{35}S]$-GTPKS binding in membrane fractions prepared from transfected mammalian cells. This assay could also be performed using baculoviruses containing nGPCR-x proteins infected into SF9 insect cells.

The neurotransmitter which activates nGPCR-x proteins can be purified to homogeneity through successive rounds of purification using nGPCR-x proteins activation as a measurement of neurotransmitter activity. The composition of the neurotransmitter can be determined by mass spectrometry and Edman degradation if peptidergic. Neurotransmitters isolated in this manner will be bioactive materials which will alter neurotransmission in the central nervous system and will produce behavioral and biochemical changes.

Example 15 - Using nGPCR-x proteins to isolate and purify G proteins

cDNAs encoding nGPCR-x proteins are epitope–tagged at the amino terminuus end of the cDNA with the cleavable influenza-hemagglutinin signal sequence followed by the FLAG epitope (IBI, New Haven, CT). Additionally, these sequences are tagged at the carboxyl terminus with DNA encoding six histidine residues. (Amino and Carboxyl Terminal Modifications to Facilitate the Production and Purification of a G Protein-Coupled Receptor, B.K. Kobilka , *Analytical Biochemistry*, Vol. 231, No. 1, Oct 1995, pp. 269-271). The resulting sequences are cloned into a baculovirus expression vector such as pVL1392 (Invitrogen). The baculovirus expression vectors are used to infect SF-9 insect cells as described (Guan,

X. M., Kobilka, T. S., and Kobilka, B. K. (1992) *J. Biol. Chem.* **267,** 21995-21998). Infected SF-9 cells could be grown in 1000-ml cultures in SF900 II medium (Life Technologies, Inc.) containing 5% fetal calf serum (Gemini, Calabasas, CA) and 0.1 mg/ml gentamicin (Life Technologies, Inc.) for 48 hours at which time the cells could be harvested. Cell membrane preparations could be separated from soluble proteins following cell lysis. nGPCR-x protein purification is carried out as described for purification of the ß2 receptor (Kobilka, Anal. Biochem., 231 (1): 269-271, 1995) including solubilization of the membranes in 0.8-1.0 % *n*-dodecyl -D-maltoside (DM) (CalBiochem, La Jolla, CA) in buffer containing protease inhibitors followed by Ni-column chromatography using chelating Sepharose™ (Pharmacia, Uppsala, Sweden). The eluate from the Ni-column is further purified on an M1 anti-FLAG antibody column (IBI). Receptor containing fractions are monitored by using receptor specific antibodies following western blot analysis or by SDS-PAGE analysis to look for an appropriate sized protein band (appropriate size would be the predicted molecular weight of the protein).

This method of purifying G protein is particularly useful to isolate G proteins that bind to the nGPCR-x proteins in the absence of an activating ligand.

Some of the preferred embodiments of the invention described above are outlined below and include, but are not limited to, the following embodiments. As those skilled in the art will appreciate, numerous changes and modifications may be made to the preferred embodiments of the invention without departing from the spirit of the invention. It is intended that all such variations fall within the scope of the invention.

The entire disclosure of each publication cited herein is hereby incorporated by reference.

What is claimed is:

1. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to even numbered sequences selected from the group consisting of: SEQ ID NO:2 to SEQ ID NO:94, SEQ ID NO:186, and fragments thereof; said nucleic acid molecule encoding at least a portion of nGPCR-x.

2. The isolated nucleic acid molecule of claim 1 comprising a sequence that encodes a polypeptide comprising even numbered sequences selected from the group consisting of SEQ ID NO:2 to SEQ ID NO:94, SEQ ID NO:186, and fragments thereof.

3. The isolated nucleic acid molecule of claim 1 comprising a sequence homologous to odd numbered sequences selected from the group consisting of SEQ ID NO:1 to SEQ ID NO:93, SEQ ID NO:185 and fragments thereof.

4. The isolated nucleic acid molecule of claim 1 comprising a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO: 93, SEQ ID NO:185 and fragments thereof.

5. The isolated nucleic acid molecule of claim 4 comprising a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.

6. The isolated nucleic acid molecule of claim 4 wherein said nucleotide sequence is selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:73, SEQ ID NO:9, SEQ ID NO:77, SEQ ID NO:11, SEQ ID NO:79, SEQ ID NO:21, SEQ ID NO:81 SEQ ID NO:53, SEQ ID NO:83, SEQ ID NO:59, SEQ ID NO:85, SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO:3, and SEQ ID NO:185.

7. The isolated nucleic acid molecule of claim 4 wherein said nucleotide sequence is selected from the group consisting of: SEQ ID NO:73, SEQ ID NO:77,

SEQ ID NO:79, SEQ ID NO:81 SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:89, SEQ ID NO:93 and SEQ ID NO:185.

8. The isolated nucleic acid molecule of claim 1 wherein said nucleic acid molecule is DNA.

9. The isolated nucleic acid molecule of claim 1 wherein said nucleic acid molecule is RNA.

10. An expression vector comprising a nucleic acid molecule of any one of claims 1 to 5.

11. The expression vector of claim 10 wherein said nucleic acid molecule comprises a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.

12. The expression vector of claim 10 wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:73, SEQ ID NO:9, SEQ ID NO:77, SEQ ID NO:11, SEQ ID NO:79, SEQ ID NO: 21, SEQ ID NO:81 SEQ ID NO:53, SEQ ID NO:83, SEQ ID NO:59, SEQ ID NO:85, SEQ ID NO:63, SEQ ID NO:87, SEQ ID NO:89, SEQ ID NO:67, SEQ ID NO:91, SEQ ID NO:93, SEQ ID NO: 3, and SEQ ID NO: 185.

13. The expression vector of claim 10 wherein said nucleotide sequence is selected from the group consisting of: SEQ ID NO: 73, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81 SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:89, SEQ ID NO:93 and SEQ ID NO: 185.

14. The expression vector of claim 10 wherein said vector is a plasmid.

15. The expression vector of claim 10 wherein said vector is a viral particle.

16. The expression vector of claim 15 wherein said vector is selected from the group consisting of adenoviruses, baculoviruses, parvoviruses, herpesviruses,

poxviruses, adeno-associated viruses, Semliki Forest viruses, vaccinia viruses, and retroviruses.

17. The expression vector of claim 10 wherein said nucleic acid molecule is operably connected to a promoter selected from the group consisting of simian virus 40, mouse mammary tumor virus, long terminal repeat of human immunodeficiency virus, maloney virus, cytomegalovirus immediate early promoter, Epstein Barr virus, rous sarcoma virus, human actin, human myosin, human hemoglobin, human muscle creatine, and human metalothionein.

18. A host cell transformed with an expression vector of claim 10.

19. The transformed host cell of claim 18 wherein said cell is a bacterial cell.

20. The transformed host cell of claim 19 wherein said bacterial cell is E. coli.

21. The transformed host cell of claim 18 wherein said cell is yeast.

22. The transformed host cell of claim 21 wherein said yeast is S. cerevisiae.

23. The transformed host cell of claim 18 wherein said cell is an insect cell.

24. The transformed host cell of claim 23 wherein said insect cell is S. frugiperda.

25. The transformed host cell of claim 18 wherein said cell is a mammalian cell.

26. The transformed host cell of claim 25 wherein mammalian cell is selected from the group consisting of chinese hamster ovary cells, HeLa cells, African green monkey kidney cells, human 293 cells, and murine 3T3 fibroblasts.

27. An isolated nucleic acid molecule comprising a nucleotide sequence complementary to at least a portion of a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185, said portion comprising at least 10 nucleotides.

28. The nucleic acid molecule of claim 27 wherein said molecule is an antisense oligonucleotide directed to a region of a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO: 185.

29. The nucleic acid molecule of claim 28 wherein said oligonucleotide is directed to a regulatory region of a sequence selected from the group of odd numbered sequences consisting of SEQ ID NO:1 to SEQ ID NO:93 and SEQ ID NO:185.

30. The nucleic acid molecule of claim 27 wherein said molecule is an antisense oligonucleotide directed to a region of nucleotide sequence selected from the group consisting of: SEQ ID NO: 73, SEQ ID NO:77, SEQ ID NO:79, SEQ ID NO:81 SEQ ID NO:83, SEQ ID NO:85, SEQ ID NO:89, SEQ ID NO:93 and SEQ ID NO: 185.

31. A composition comprising a nucleic acid molecule of any one of claims 1 to 5 or 27 and an acceptable carrier or diluent.

32. A composition comprising a recombinant expression vector of claim 10 and an acceptable carrier or diluent.

33. A method of producing a polypeptide that comprises a sequence selected from the group of even numbered sequences consisting SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186, and homologs and fragments thereof, said method comprising the steps of:

a) introducing a recombinant expression vector of claim 10 into a compatible host cell;

b) growing said host cell under conditions for expression of said polypeptide; and

c) recovering said polypeptide.

34. The method of claim 33 wherein said host cell is lysed and said polypeptide is recovered from the lysate of said host cell.

35. The method of claim 33 wherein said polypeptide is recovered by purifying the culture medium without lysing said host cell.

36. An isolated polypeptide encoded by a nucleic acid molecule of claim 1.

37. The polypeptide of claim 36 wherein said polypeptide comprises a sequence selected from the group of even numbered sequences consisting SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.

38. The polypeptide of claim 36 wherein said polypeptide comprises an amino acid sequence homologous to a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.

39. The polypeptide of claim 36 wherein said sequence homologous to a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186 comprises at least one conservative amino acid substitution compared to the even numbered sequences in the group of even numbered sequences consisting of SEQ ID NO: 2 to SEQ ID NO: 94 and SEQ ID NO: 186.

40. The polypeptide of claim 36 wherein said polypeptide comprises a fragment of a polypeptide with a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.

41. The polypeptide of claim 36 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.

42. The polypeptide of claim 36 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186; SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.

43. A composition comprising a polypeptide of claim 36 and an acceptable carrier or diluent.

44. An isolated antibody which binds to an epitope on a polypeptide of claim 36.

45. The antibody of claim 44 wherein said antibody is a monoclonal antibody.

46. A composition comprising an antibody of claim 44 and an acceptable carrier or diluent.

47. A method of inducing an immune response in a mammal against a polypeptide of claim 36 comprising administering to said mammal an amount of said polypeptide sufficient to induce said immune response.

48. A method for identifying a compound which binds nGPCR-x comprising the steps of:

a) contacting nGPCR-x with a compound; and

b) determining whether said compound binds nGPCR-x.

49. The method of claim 48 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.

50. The method of claim 48 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186; SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.

51. The method of claim 48 wherein binding of said compound to nGPCR-x is determined by a protein binding assay.

52. The method of claim 48 wherein said protein binding assay is selected from the group consisting of a gel-shift assay, Western blot, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, and ELISA.

53. A compound identified by the method of claim 48.

54. A method for identifying a compound which binds a nucleic acid molecule encoding nGPCR-x comprising the steps of:

a) contacting said nucleic acid molecule encoding nGPCR-x with a compound; and

b) determining whether said compound binds said nucleic acid molecule.

55. The method of claim 54 wherein binding is determined by a gel-shift assay.

56. A compound identified by the method of claim 54.

57. A method for identifying a compound which modulates the activity of nGPCR-x comprising the steps of:

a) contacting nGPCR-x with a compound; and

b) determining whether nGPCR-x activity has been modulated.

58. The method of claim 57 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.

59. The method of claim 57 wherein the nGPCR-x comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186;

SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.

60. The method of claim 57 wherein said activity is neuropeptide binding.

61. The method of claim 57 wherein said activity is neuropeptide signaling.

62. A compound identified by the method of claim 57.

63. A method of identifying an animal homolog of nGPCR-x comprising the steps:

a) comparing the nucleic acid sequences of the animal with a sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof, said portions being at least 10 nucleotides; and

b) identifying nucleic acid sequences of the animal that are homologous to said sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof.

64. The method of claim 63 wherein comparing the nucleic acid sequences of the animal with a sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof, said portions being at least 10 nucleotides is performed by DNA hybridization.

65. The method of claim 63 wherein comparing the nucleic acid sequences of the animal with a sequence selected from the group of odd numbered sequence consisting of SEQ ID NO: 1 to SEQ ID NO: 93, SEQ ID NO: 185, and portions thereof, said portions being at least 10 nucleotides is performed by computer homology search.

66. A method of screening a human subject to diagnose a disorder affecting the brain or genetic predisposition therefor, comprising the steps of:

(a) assaying nucleic acid of a human subject to determine a presence or an absence of a mutation altering an amino acid sequence, expression, or biological

activity of at least one nGPCR that is expressed in the brain, wherein the nGPCR comprises an amino acid sequence selected from the group consisting of: SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof, and wherein the nucleic acid corresponds to a gene encoding the nGPCR; and

(b) diagnosing the disorder or predisposition from the presence or absence of said mutation, wherein the presence of a mutation altering the amino acid sequence, expression, or biological activity of the nGPCR in the nucleic acid correlates with an increased risk of developing the disorder.

67. A method according to claim 66, wherein the nGPCR is nGPCR-40 comprising an amino acid sequence set forth in SEQ ID NO:84 or an allelic variant thereof.

68. A method according to claim 66, wherein the nGPCR is nGPCR-54 comprising an amino acid sequence set forth in SEQ ID NO:86 or an allelic variant thereof.

69. A method according to claim 66, wherein the disease is schizophrenia.

70. A method according to claim 66, wherein the assaying step comprises at least one procedure selected from the group consisting of:

a) comparing nucleotide sequences from the human subject and reference sequences and determining a difference of either

at least a nucleotide of at least one codon between the nucleotide sequences from the human subject that encodes an nGPCR-40 allele and an nGPCR-40 reference sequence, or

at least a nucleotide of at least one codon between the nucleotide sequences from the human subject that encodes an nGPCR-54 allele and an nGPCR-54 reference sequence;

(b) performing a hybridization assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences;

(c) performing a polynucleotide migration assay to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences; and

(d) performing a restriction endonuclease digestion to determine whether nucleic acid from the human subject has a nucleotide sequence identical to or different from one or more reference sequences.

71. A method according to claim 70 wherein the assaying step comprises: performing a polymerase chain reaction assay to amplify nucleic acid comprising nGPCR-40 or nGPCR-54 coding sequence, and determining nucleotide sequence of the amplified nucleic acid.

72. A method of screening for an nGPCR-40 or nGPCR-54 hereditary schizophrenia genotype in a human patient, comprising the steps of:

(a) providing a biological sample comprising nucleic acid from said patient, said nucleic acid including sequences corresponding to alleles of nGPCR-40 or nGPCR-54; and

(b) detecting the presence of one or more mutations in the nGPCR-40 allele or the nGPCR-54 allele;

wherein the presence of a mutation in an nGPCR-40 allele or nGPCR-54 allele is indicative of a hereditary schizophrenia genotype.

73. The method according to claim 72 wherein said biological sample is a cell sample.

74. The method according to claim 72 wherein said detecting the presence of a mutation comprises sequencing at least a portion of said nucleic acid, said portion comprising at least one codon of said nGPCR-40 or nGPCR-54 alleles.

75. The method according to claim 72 wherein said nucleic acid is DNA.

76. The method according to claim 72 wherein said nucleic acid is RNA.

77. A kit for screening a human subject to diagnose schizophrenia or a genetic predisposition therefor, comprising, in association:

(a) an oligonucleotide useful as a probe for identifying polymorphisms in a human nGPCR-40 gene or a human nGPCR-54 gene, the oligonucleotide comprising 6-50 nucleotides in a sequence that is identical or complementary to a sequence of a wild type human nGPCR-40 or nGPCR-54 gene sequence or nGPCR-40 or nGPCR-54 coding sequence, except for one sequence difference selected from the group consisting of a nucleotide addition, a nucleotide deletion, or nucleotide substitution; and

(b) a media packaged with the oligonucleotide, said media containing information for identifying polymorphisms that correlate with schizophrenia or a genetic predisposition therefor, the polymophisms being identifiable using the oligonucleotide as a probe.

78. A method of identifying a nGPCR allelic variant that correlates with a mental disorder, comprising steps of:

(a) providing a biological sample comprising nucleic acid from a human patient diagnosed with a mental disorder, or from the patient's genetic progenitors or progeny;

(b) detecting in the nucleic acid the presence of one or more mutations in an nGPCR that is expressed in the brain, wherein the nGPCR comprises an amino acid sequence selected from the group consisting of SEQ ID NO:74, SEQ ID NO:186, SEQ ID NO:78, SEQ ID NO:80, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:86, SEQ ID NO:90, and SEQ ID NO:94, and allelic variants thereof, and wherein the nucleic acid includes sequence corresponding to the gene or genes encoding nGPCR;

wherein the one or more mutations detected indicates an allelic variant that correlates with a mental disorder.

79. A method according to claim 78, wherein the disorder is schizophrenia, and wherein the at least one nGPCR is nGPCR-40, nGPCR-54, or an allelic variant thereof.

80. A purified and isolated polynucleotide comprising a nucleotide sequence encoding an nGPCR-40 or nGPCR-54 allelic variant identified according to claim 79.

81. A host cell transformed or transfected with a polynucleotide according to claim 80 or with a vector comprising the polynucleotide.

82. A purified polynucleotide comprising a nucleotide sequence encoding nGPCR-40 or nGPCR-54 of a human with schizophrenia;

wherein said polynucleotide hybridizes to the complement of SEQ ID NO:83 or of SEQ ID NO:85 under the following hybridization conditions:

(a) hybridization for 16 hours at 42°C in a hybridization solution comprising 50% formamide, 1% SDS, 1 M NaC1, 10% dextran sulfate and

(b) washing 2 times for 30 minutes at 60°C in a wash solution comprising 0.1x SSC and 1% SDS; and

wherein the polynucleotide that encodes nGPCR-40 or nGPCR-54 amino acid sequence of the human differs from SEQ ID NO:84 or SEQ ID NO:86 by at least one residue.

83. A vector comprising a polynucleotide according to claim 82.

84. A host cell that has been transformed or transfected with a polynucleotide according to claim 82 and that expresses the nGPCR-40 or nGPCR-54 protein encoded by the polynucleotide.

85. A host cell according to claim 84 that has been co-transfected with a polynucleotide encoding the nGPCR-40 or nGPCR-54 amino acid sequence set forth in SEQ ID NO:84 or SEQ ID NO:86 and that expresses the nGPCR-40 or nGPCR-54 protein having the amino acid sequence set forth in SEQ ID NO:84 or SEQ ID NO:86.

86. A method for identifying a modulator of biological activity of nGPCR-40 or nGPCR-54 comprising the steps of:

a) contacting a cell according to claim 84 in the presence and in the absence of a putative modulator compound;

b) measuring nGPCR-40 or nGPCR-54 biological activity in the cell;

wherein decreased or increased nGPCR-40 or nGPCR-54 biological activity in the presence versus absence of the putative modulator is indicative of a modulator of biological activity.

87. A method to identify compounds useful for the treatment of schizophrenia, said method comprising steps of:

(a) contacting a composition comprising nGPCR-40 with a compound suspected of binding nGPCR-40 or contacting a composition comprising nGPCR-54 with a compound suspected of binding nGPCR-54;

(b) detecting binding between nGPCR-40 and the compound suspected of binding nGPCR-40 or between nGPCR-54 and the compound suspected of binding nGPCR-54;

wherein compounds identified as binding nGPCR-40 or nGPCR-54 are candidate compounds useful for the treatment of schizophrenia.

88. A method for identifying a compound useful as a modulator of binding between nGPCR-40 and a binding partner of nGPCR-40 or between nGPCR-54 and a binding partner of nGPCR-54 comprising the steps of:

(a) contacting the binding partner and a composition comprising nGPCR-40 or nGPCR-54 in the presence and in the absence of a putative modulator compound;

(b) detecting binding between the binding partner and nGPCR-40 or nGPCR-54;

wherein decreased or increased binding between the binding partner and nGPCR-40 or nGPCR-54 in the presence of the putative modulator, as compared to binding in the absence of the putative modulator is indicative a modulator compound useful for the treatment of schizophrenia.

89. A method according to claim 87 or 88 wherein the composition comprises a cell expressing nGPCR-40 or nGPCR-54 on its surface.

90. An method according to claim 89 wherein the composition comprises a cell transformed or transfected with a polynucleotide that encodes nGPCR-40 or nGPCR-54.

91. A method of purifying a G protein from a sample containing said G protein comprising the steps of:

a) contacting said sample with a polypeptide of claim 1 for a time sufficient to allow said G protein to form a complex with said polypeptide;

b) isolating said complex from remaining components of said sample;

c) maintaining said complex under conditions which result in dissociation of said G protein from said polypeptide; and

d) isolating said G protein from said polypeptide.

92. The method of claim 91 wherein said sample comprises an amino acid sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.

93. The method of claim 91 wherein said polypeptide comprises an amino acid sequence homologous to a sequence selected from the group of even numbered sequences consisting of SEQ ID NO:2 to SEQ ID NO:94 and SEQ ID NO:186.

94. The method of claim 91 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 2, SEQ ID NO: 74; SEQ ID NO: 4, SEQ ID NO: 186; SEQ ID NO:10, SEQ ID NO:78; SEQ ID NO:12, SEQ ID NO:80; SEQ ID NO: 22, SEQ ID NO:82; SEQ ID NO:54, SEQ ID NO:84; SEQ ID NO:60, SEQ ID NO: 86; SEQ ID NO:64, SEQ ID NO: 88, SEQ ID NO:90; SEQ ID NO:68, SEQ ID NO: 92, and SEQ ID NO:94.

95. The method of claim 91 wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 74; SEQ ID NO: 186; SEQ ID NO:78; SEQ ID NO:80; SEQ ID NO:82; SEQ ID NO:84; SEQ ID NO: 86; SEQ ID NO:90; and SEQ ID NO:94.

96. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence homologous to SEQ ID NO:76, and fragments thereof; said nucleic acid molecule encoding at least a portion of nGPCR-5.

97. An isolated polypeptide encoded by a nucleic acid molecule of claim 96.

SEQUENCE LISTING

<110> Pharmacia & Upjohn Company
Vogeli, Gabriel
Huff, Rita
Sejlitz, Torsten
Lind, Peter
Slightom, Jerry
Schellin, Kathleen
Bannigan, Chris
Ruff, Valerie
Kaytes, Paul
Wood, Linda
Parodi, Luis
Hiebsch, Ronald

<120> Novel G Protein Coupled Receptors

<130> 043P1FHRM296

<150> 60/165,838
<151> 1999-11-16

<150> 60/198,568
<151> 2000-04-20

<150> 60/166,071
<151> 1999-11-17

<150> 60/166,678
<151> 1999-11-19

<150> 60/173,396
<151> 1999-12-28

<150> 60/184,129
<151> 2000-02-22

<150> 60/185,421
<151> 2000-02-28

<150> 60/185,554
<151> 2000-02-28

<150> 60/186,530
<151> 2000-03-02

<150> 60/186,811
<151> 2000-03-03

<150> 60/188,114
<151> 2000-03-09

<150> 60/190,310
<151> 2000-03-17

<150> 60/190,800
<151> 2000-03-21

<150> 60/201,190
<151> 2000-05-02

<150> 60/203,111
<151> 2000-05-08

<150> 60/207,094
<151> 2000-05-25

<160> 190

<170> PatentIn version 3.0

<210> 1
<211> 1182
<212> DNA
<213> H.Sapiens

<400> 1
gtctgggggt gggggatgct gggacagggg tcaattgcct gaagcaagtg ctctcatccc 60

cctagctcct gctgatctag ttggggctcc agagtgggga ggagaaaggc actttgaaac 120

ttctctgccc ttaccgtctt agccatcaaa ctctgagctg gagatagtga cgatgtgaca 180

ggaactttcc ctgggcctct ctgggccaca attcctggcc gagagaaaga ggaggaatga 240

ggtgagcacc ttcttcactc ctagggccat gtggtagagc tgcagtcgca cctccttctg 300

ccaataggca tagatgagtg ggttgagcag ggagttgccc acgccgagca gccacaggta 360

ccgttccagc actaggtaga ggtgacactc ctggcaggcc acctgcacaa tgccagtgat 420

aaggaagggg gtccaggata gagcaaagct cccaatgaga acagacacag tacggagagc 480

tttgaagtcg ctgggagtcc gtggggatcg ataacctcca gccatggctc ctgcatgttc 540

catctttcga atctgctggc tgtgcatgga ggcaatcttg agcatgtcgc agtagaagaa 600

gacaaagagg agcatggctg ggaagaagcc aacgcaggag agggtcagca cgaagtgagg 660

gtgaaataca gcaaagaagc tgcactgccc tttgtaggca gtctgctgga acatggggat 720

tccgagtggg aggaagccaa tgaggtaaga cactaaccac agcccggcaa tgcaggcccc 780

ggccacgaac ccactcatga tcttcaagta gcggaagggc tgcttgatgg caaggtacct 840

gtcaaaggtg atcagcatga ccgtgaggac agaggcagct gcggaggaag tgacaaatgc 900

catccgcagg ctgcacaggg tcttctgtgt gggccgagaa gggctggaga gctggtctgt 960

gagtaggcca gagatggcca caccaatcaa ggtgtcagcc acagccagat tcaaggtgaa 1020

gcagagactg acaccatcat tcttgtggat caacagcagc acagccacag ccactagtgt 1080

gttagtagca atgatgaggg aggccaggac agcaaggatc actccaaatg agaaagatga 1140

ttccatgtct cgaagtggca ggacttcact taccagggca tg 1182

<210> 2
<211> 335
<212> PRT
<213> H.Sapiens

<400> 2

Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser
1 5 10 15

Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu
20 25 30

Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala
35 40 45

Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp
50 55 60

Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu
65 70 75 80

Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val
85 90 95

Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg
100 105 110

Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly
115 120 125

Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro
130 135 140

Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val
145 150 155 160

Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro
165 170 175

Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala
180 185 190

Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met
195 200 205

Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu
210 215 220

Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro
225 230 235 240

Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu
245 250 255

Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser
260 265 270

Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu
275 280 285

Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe
290 295 300

Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu
305 310 315 320

Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly
325 330 335

<210> 3
<211> 657
<212> DNA
<213> H.Sapiens

<400> 3
cagcgcgagc gccttcatgg tgacggtgtc catgcgctgg cagtgtctgc gtgccacccg 60

gtgcacctgg agcgaggtga ggcagagcac cgccagcggc agcacgaagc ccacggcatg 120

gagcgtggcg gtgaaggctg cgaagcgcgg acgctcaggc tcgggcggca ggcgcagcga 180

acaggacgcg aaggcgctgc tgtagccaag ccacgagcag ccaagtgcag cgcctgagaa 240

ggccagcgac tgtccccagg cacagcccag cagcaggccg gcatagcgcg gtcgcaggcg 300

tccggcgtag cgcagtggga agcccactgc cagccactgg tctgcgctca gcgccgccac 360

gctcagcgcc gcgttggacg ccaggaaggt gtccaggaag ccaatgactt ggcatgcgcc 420

gggcgccgac ggtgtccgcc cgcgcatcac accgagcagc gtgaagggca tgtccagcgc 480

cgccagcagc aggtggccca gagacagatt caccaggagg acgcctgagg ctcgagtgcg 540

gagctcagcg ctgtaggcgc aacaaagcag caccagtgcg ttggatagca gcgccacggc 600

cagtaccatc accaggagac ccgccagcag cgcctcgccg gggcccatgg cgctagc 657

<210> 4
<211> 217
<212> PRT
<213> H.Sapiens

<400> 4

Ser Ala Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met
1 5 10 15

Val Leu Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys
20 25 30

Ala Tyr Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val
35 40 45

Asn Leu Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe
50 55 60

Thr Leu Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala
65 70 75 80

Cys Gln Val Ile Gly Phe Leu Asp Thr Phe Leu Ala Ser Asn Ala Ala
85 90 95

Leu Ser Val Ala Ala Leu Ser Ala Asp Gln Trp Leu Ala Val Gly Phe
100 105 110

Pro Leu Arg Tyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu
115 120 125

Leu Gly Cys Ala Trp Gly Gln Ser Leu Ala Phe Ser Gly Ala Ala Leu
130 135 140

Gly Cys Ser Trp Leu Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu
145 150 155 160

Arg Leu Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Phe Thr Ala
165 170 175

Thr Leu His Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cys Leu
180 185 190

Thr Ser Leu Gln Val His Arg Val Ala Arg Arg His Cys Gln Arg Met
195 200 205

Asp Thr Val Thr Met Lys Ala Leu Ala
210 215

<210> 5
<211> 222
<212> DNA
<213> H.Sapiens

<400> 5
tgtgcaggtg tgatctccat tcctttgtac atccctcaca cgctgttcga atgggatttt 60

ggaaaggaaa tctgtgtatt ttggctcact actgactatc tgttatgtac agcatctgta 120

tataacattg tcctcatcag ctatgatcga tacctgtcag tctcaaatgc tgtaagtcga 180

acacattaat ttatccccct tagaagatta tgtaaatgta ta 222

<210> 6
<211> 73
<212> PRT
<213> H.Sapiens

<400> 6

Cys Ala Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe
1 5 10 15

Glu Trp Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp
20 25 30

Tyr Leu Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr
35 40 45

Asp Arg Tyr Leu Ser Val Ser Asn Ala Val Ser Arg Thr His Phe Ile
50 55 60

Pro Leu Arg Arg Leu Cys Lys Cys Ile
65 70

<210> 7
<211> 507
<212> DNA
<213> H.Sapiens

<400> 7
gacgtcgaag caggtgatga tgcccagggc gtgcaccggg taggtgagat cggtgcgcgc 60

cagcggggac agggcggtca ggagcagcag ccaggtccct gcacacgcgg ccaccgcgta 120

acgacggcgg cgccagcgct tggagctgag cgggtacagg atccccagga agcgctccac 180

gctgatacag gtcatggtga ggatgctgga atacatgttt gcgtaaaagg ccacggtcac 240

cacgttgcaa agcagcaccc cgaataccca gtggtggcgg ttgcaatggt agtagatttg 300

gaaaggcaac acgctggcca gcatcaggtc cgtgacgctc aggttgatca tgaagatgac 360

cgacggggat ctgggcccca tgcgccggca cagcacccac agagagaaga ggttgcccgg 420

gatgctgacc gccgccacca gcgagtacac cacgggcagg gccaccgcga tcgccgggtt 480

ccgcagcatc tgcagcgtcg cgttgtc 507

<210> 8
<211> 169
<212> PRT
<213> H.Sapiens

<400> 8

Asp Asn Ala Thr Leu Gln Met Leu Arg Asn Pro Ala Ile Ala Val Ala
1 5 10 15

Leu Pro Val Val Tyr Ser Leu Val Ala Ala Val Ser Ile Pro Gly Asn
20 25 30

Leu Phe Ser Leu Trp Val Leu Cys Arg Arg Met Gly Pro Arg Ser Pro
35 40 45

Ser Val Ile Phe Met Ile Asn Leu Ser Val Thr Asp Leu Met Leu Ala
50 55 60

Ser Val Leu Pro Phe Gln Ile Tyr Tyr His Cys Asn Arg His His Trp
65 70 75 80

Val Phe Gly Val Leu Cys Asn Leu Val Val Thr Val Ala Phe Tyr Ala
85 90 95

Asn Met Tyr Ser Ser Ile Leu Thr Met Thr Cys Ile Ser Val Glu Arg
100 105 110

Phe Leu Gly Ile Leu Tyr Pro Leu Ser Ser Lys Arg Trp Arg Arg Arg
115 120 125

Arg Tyr Ala Val Ala Ala Cys Ala Gly Thr Trp Leu Leu Leu Leu Thr
130 135 140

Ala Leu Ser Pro Leu Ala Arg Thr Asp Leu Thr Tyr Pro Val His Ala
145 150 155 160

Leu Gly Ile Ile Thr Cys Phe Asp Val
165

<210> 9
<211> 270
<212> DNA
<213> H.Sapiens

<400> 9
cccatgttcc tgctcctggg cagcctcacg ttgtcggatc tgctggcagg cgccgcctac 60

gccgccaaca tcctactgtc ggggccgctc acgctgaaac tgtccccgc gctctggttc 120

gcacgggagg gaggcgtctt cgtggcactc actgcgtccg tgctgagcct cctgggcatc 180

gcgctggagc gcagcctcac catggcgcgc agggggcccg cgcccgtctc cagtcggggg 240

cgcacgctgg cgatggcagc cgcggcctgg 270

<210> 10
<211> 90
<212> PRT
<213> H.Sapiens

<400> 10

Pro Met Phe Leu Leu Leu Gly Ser Leu Thr Leu Ser Asp Leu Leu Ala
1 5 10 15

Gly Ala Ala Tyr Ala Ala Asn Ile Leu Leu Ser Gly Pro Leu Thr Leu
20 25 30

Lys Leu Ser Pro Ala Leu Trp Phe Ala Arg Glu Gly Gly Val Phe Val
35 40 45

Ala Leu Thr Ala Ser Val Leu Ser Leu Leu Gly Ile Ala Leu Glu Arg
50 55 60

Ser Leu Thr Met Ala Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Gly
65 70 75 80

Arg Thr Leu Ala Met Ala Ala Ala Ala Trp
85 90

<210> 11
<211> 888

<212> DNA
<213> H.Sapiens

<400> 11
```
ctgctcattg tggcctttgt gctgggcgca ctaggcaatg gggtcgccct gtgtggtttc     60

tgcttccaca tgaagacctg gaagcccagc actgtttacc ttttcaattt ggccgtggct    120

gatttcctcc ttatgatctg cctgcctttt cggacagact attacctcag acgtagacac    180

tgggcttttg gggacattcc ctgccgagtg gggctcttca cgttggccat gaacagggcc    240

gggagcatcg tgttccttac ggtggtggct gcggacaggt atttcaaagt ggtccacccc    300

caccacgcgg tgaacactat ctccacccgg gtggcggctg gcatcgtctg caccctgtgg    360

gccctggtca tcctgggaac agtgtatctt ttgctggaga accatctctg cgtgcaagag    420

acggccgtct cctgtgagag cttcatcatg gagtcggcca atggctggca tgacatcatg    480

ttccagctgg agttctttat gcccctcggc atcatcttat tttgctcctt caagattgtt    540

tggagcctga ggcggaggca gcagctggcc agacaggctc ggatgaagaa ggcgacccgg    600

ttcatcatgg tggtggcaat tgtgttcatc acatgctacc tgcccagcgt gtctgctaga    660

ctctatttcc tctggacggt gccctcgagt gcctgcgatc cctctgtcca tggggccctg    720

cacataaccc tcagcttcac ctacatgaac agcatgctgg atcccctggt gtattatttt    780

tcaagcccct cctttcccaa attctacaac aagctcaaaa tctgcagtct gaaacccaag    840

cagccaggac actcaaaaac acaaaggccg gaagagatgc caatttcg                 888
```

<210> 12
<211> 296
<212> PRT
<213> H.Sapiens

<400> 12

```
Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly Asn Gly Val Ala
1               5                   10                  15

Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys Pro Ser Thr Val
            20                  25                  30

Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu Met Ile Cys Leu
        35                  40                  45

Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His Trp Ala Phe Gly
    50                  55                  60

Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala Met Asn Arg Ala
65                  70                  75                  80

Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp Arg Tyr Phe Lys
                85                  90                  95
```

```
Val Val His Pro His His Ala Val Asn Thr Ile Ser Thr Arg Val Ala
            100                 105                 110

Ala Gly Ile Val Cys Thr Leu Trp Ala Leu Val Ile Leu Gly Thr Val
        115                 120                 125

Tyr Leu Leu Leu Glu Asn His Leu Cys Val Gln Glu Thr Ala Val Ser
    130                 135                 140

Cys Glu Ser Phe Ile Met Glu Ser Ala Asn Gly Trp His Asp Ile Met
145                 150                 155                 160

Phe Gln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile Leu Phe Cys Ser
                165                 170                 175

Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln Leu Ala Arg Gln
            180                 185                 190

Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val Val Ala Ile Val
        195                 200                 205

Phe Ile Thr Cys Tyr Leu Pro Ser Val Ser Ala Arg Leu Tyr Phe Leu
    210                 215                 220

Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val His Gly Ala Leu
225                 230                 235                 240

His Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met Leu Asp Pro Leu
                245                 250                 255

Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe Tyr Asn Lys Leu
            260                 265                 270

Lys Ile Cys Ser Leu Lys Pro Lys Gln Pro Gly His Ser Lys Thr Gln
        275                 280                 285

Arg Pro Glu Glu Met Pro Ile Ser
    290                 295

&lt;210&gt;  13
&lt;211&gt;  510
&lt;212&gt;  DNA
&lt;213&gt;  H.Sapiens

&lt;400&gt;  13
tggagctgtg ccaccaccta tctggtgaac ctgatggtgg ccgacctgct ttatgtgcta      60

ttgcccttcc tcatcatcac ctactcacta gatgacaggt ggcccttcgg ggagctgctc     120

tgcaagctgg tgcacttcct gttctatatc aacctttacg gcagcatcct gctgctgacc     180

tgcatctctg tgcaccagtt cctaggtgtg tgccacccac tgtgttcgct gccctaccgg     240

acccgcaggc atgcctggct gggcaccagc accacctggg ccctggtggt cctccagctg     300

ctgcccacac tggccttctc ccacacggac tacatcaatg gccagatgat ctggtatgac     360

atgaccagcc aagagaattt tgatcggctt tttgcctacg gcatagttct gacattgtct     420
```

ggctttcttt ccctccttgg tcattttggt gtgctattca ctgatggtca ggagcctgat 480

caagccagag gagaacctca tgaggacagg 510

<210> 14
<211> 170
<212> PRT
<213> H.Sapiens

<400> 14

Trp Ser Cys Ala Thr Thr Tyr Leu Val Asn Leu Met Val Ala Asp Leu
1 5 10 15

Leu Tyr Val Leu Leu Pro Phe Leu Ile Ile Thr Tyr Ser Leu Asp Asp
20 25 30

Arg Trp Pro Phe Gly Glu Leu Leu Cys Lys Leu Val His Phe Leu Phe
35 40 45

Tyr Ile Asn Leu Tyr Gly Ser Ile Leu Leu Leu Thr Cys Ile Ser Val
50 55 60

His Gln Phe Leu Gly Val Cys His Pro Leu Cys Ser Leu Pro Tyr Arg
65 70 75 80

Thr Arg Arg His Ala Trp Leu Gly Thr Ser Thr Thr Trp Ala Leu Val
85 90 95

Val Leu Gln Leu Leu Pro Thr Leu Ala Phe Ser His Thr Asp Tyr Ile
100 105 110

Asn Gly Gln Met Ile Trp Tyr Asp Met Thr Ser Gln Glu Asn Phe Asp
115 120 125

Arg Leu Phe Ala Tyr Gly Ile Val Leu Thr Leu Ser Gly Phe Leu Ser
130 135 140

Leu Leu Gly His Phe Gly Val Leu Phe Thr Asp Gly Gln Glu Pro Asp
145 150 155 160

Gln Ala Arg Gly Glu Pro His Glu Asp Arg
165 170

<210> 15
<211> 894
<212> DNA
<213> H.Sapiens

<220>
<221> misc_feature
<222> (431)..(461)
<223> n is any nucleotide

<400> 15
ccaccacgcg cagcacgccg acagggcctc tccctcccat tctcccgcag gcccggacga 60

ccacgctgcc tccagccggt cggcaaacta gggcagctcg cagcccacga acagcagccc 120

cagcagctgg ctcatcttca ggctctgcac cttggcgcgg ggcatcgcgc tgggcgcacg 180

ggctccacct gggctcgccg accaggccgc tgcacccgct ggggccttca gccggtgccg 240

ccaccagacg gagagtaggt ggccacaagc gacacccatg atcttaacag gcgcgacgaa 300

gcccgcgacg gcctcataga acgcgtacac ctgcacgtgc cagcgctgca ggagcgcgaa 360

gatccagtgg cagcgacgca tccccggcca ggctcggcgg gagagtggcg cgcctggctg 420

cagagacgtt nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nagtactagc gcaccacaaa 480

ccccgacccc cgcgccagca gcagtgccag cagccagccc agggcggcga gggcacgcgc 540

gggcagcggc cggccgtgcg gaagacgcac cgcgcgccgg cgctcgaggg cgatgagcac 600

cacgaggtgg gccgaggcgc cccgcccgga tgcctgcagc agctgcagga agcggcacgc 660

caggtccccc gtggccgcgc ggggctcgcc cagcagttcc caggccagct gtgacagcgc 720

ccgtgccccg cacgcgtaca ggtccgccag ggccagctgc accagcagga agtccatctt 780

gcgacgcttn nnnnnnnnnn nnnnnnnnnn nnnnnnnnac aggcggcaca gcactgtggt 840

gttgcctgcc accgccacca ccaggatgac ccccaggaac accaggcgga cgcg 894

<210> 16
<211> 296
<212> PRT
<213> H.Sapiens

<220>
<221> UNSURE
<222> (26)..(35)
<223> Xaa is unknown

<220>
<221> UNSURE
<222> (144)..(154)
<223> Xaa is Unknown

<400> 16

Arg Val Arg Leu Val Phe Leu Gly Val Ile Leu Val Val Ala Val Ala
1 5 10 15

Gly Asn Thr Thr Val Leu Cys Arg Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Lys Arg Arg Lys Met Asp Phe Leu Leu Val Gln Leu Ala
35 40 45

Leu Ala Asp Leu Tyr Ala Cys Gly Gly Thr Ala Leu Ser Gln Leu Ala

50 55 60

Trp Glu Leu Leu Gly Glu Pro Arg Ala Ala Thr Gly Asp Leu Ala Cys
65 70 75 80

Arg Phe Leu Gln Leu Leu Gln Ala Ser Gly Arg Gly Ala Ser Ala His
85 90 95

Leu Val Val Leu Ile Ala Leu Glu Arg Arg Arg Ala Val Arg Leu Pro
100 105 110

His Gly Arg Pro Leu Pro Ala Arg Ala Leu Ala Ala Leu Gly Trp Leu
115 120 125

Leu Ala Leu Leu Leu Ala Arg Gly Ser Gly Phe Val Val Arg Tyr Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser Leu Gln Pro Gly
145 150 155 160

Ala Pro Leu Ser Ala Arg Ala Trp Pro Gly Met Arg Arg Cys His Trp
165 170 175

Ile Phe Ala Leu Leu Gln Arg Trp His Val Gln Val Tyr Ala Phe Tyr
180 185 190

Glu Ala Val Ala Gly Phe Val Ala Pro Val Lys Ile Met Gly Val Ala
195 200 205

Cys Gly His Leu Leu Ser Val Trp Trp Arg His Arg Leu Lys Ala Pro
210 215 220

Ala Gly Ala Ala Ala Trp Ser Ala Ser Pro Gly Gly Ala Arg Ala Pro
225 230 235 240

Ser Ala Met Pro Arg Ala Lys Val Gln Ser Leu Lys Met Ser Gln Leu
245 250 255

Leu Gly Leu Leu Phe Val Gly Cys Glu Leu Pro Phe Ala Asp Arg Leu
260 265 270

Glu Ala Ala Trp Ser Ser Gly Pro Ala Gly Glu Trp Glu Gly Glu Ala
275 280 285

Leu Ser Ala Cys Cys Ala Trp Trp
290 295

<210> 17
<211> 801
<212> DNA
<213> H.Sapiens

<400> 17
tctaagtttt tctctgaact ttgagcctgt gaaaaaagaa gggatgctgc ctcaggccac 60

cccagcctag atactcactc tgagtgccat gaggtagtag aggacactga tgacagtcat 120

ggggaggagg tagaatagga aggaggtgac ctggatgatg aaattgtaga tccacatggg 180

```
cttgatgacc gtacaggtgg ccgaacctgg gaccagggac ccattgggga agtagtggaa     240
cttgatgcca tggatgctgg tgttgggcag ggagaagagc acggagaagc cccagacgat     300
gccgaggatc ctgagggccc ggcgccgggt gctctgcagt ttggcgcgga acgggtgtag     360
gatggccacg tagcgctcca cgctgacggt ggtgatgctg aggatggagg cgaagcacac     420
ggtctcaaag agggccgtct tgaagtagca gcccacgggc ccgaacaaga aagggtagtt     480
gcgccacatc tcatagacct ccaggggcat tccaaggagc aggaccagga ggtcagagac     540
cgccaggctg aagaggtagt agttggtggg cgtcttcata gcctggtgct gcagaatcac     600
caggcacacc aggacattgc caatgacccc caccacaaaa attggcacat acaccacaga     660
cacggggagg aagaagtggc tgcgccgagg tccgcagagg aaggccagat actcctcggt     720
gctgttcagg tgtttctgga atggatcttc tagtttctgc tggtagatcc aggaagcatt     780
ctgaagtttt tccatccctg a                                               801
```

```
<210>  18
<211>  249
<212>  PRT
<213>  H.Sapiens

<400>  18

Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys
1               5                   10                  15

Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu
            20                  25                  30

Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser
        35                  40                  45

Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu
    50                  55                  60

Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Asn Thr
65                  70                  75                  80

Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu
                85                  90                  95

Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu
            100                 105                 110

Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val
        115                 120                 125

Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr
    130                 135                 140

Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg
145                 150                 155                 160
```

Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe
165 170 175

Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro
180 185 190

Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro
195 200 205

Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr
210 215 220

Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu
225 230 235 240

Arg Val Ser Ile Ala Gly Val Ala Gly
245

<210> 19
<211> 222
<212> DNA
<213> H.Sapiens

<400> 19
atcaagatga tttttgctat cgtgcaaatt attggatttt ccaactccat ctgtaatccc 60

attgtctatg catttatgaa tgaaaacttc aaaaaaaatg ttttgtctgc agtttgttat 120

tgcatagtaa ataaaacctt ctctccagca caaaggcatg gaaattcagg aattacaatg 180

atgcggaaga aagcaaagtt ttccctcaga gagaatccag tg 222

<210> 20
<211> 73
<212> PRT
<213> H.Sapiens

<400> 20

Ile Lys Met Ile Phe Ala Ile Val Gln Ile Ile Gly Phe Ser Asn Ser
1 5 10 15

Ile Cys Asn Pro Ile Val Tyr Ala Phe Met Asn Glu Asn Phe Lys Lys
20 25 30

Asn Val Leu Ser Ala Val Cys Tyr Cys Ile Val Asn Lys Thr Phe Ser
35 40 45

Pro Ala Gln Arg His Gly Asn Ser Gly Ile Thr Met Met Arg Lys Lys
50 55 60

Ala Lys Phe Ser Leu Arg Glu Asn Pro
65 70

<210> 21
<211> 447
<212> DNA

<213> H.Sapiens

<400> 21
gccacagcat gcagttttct gtagaattcc actttgtctt tgcacttgaa gaagatgagg 60

tatctggtga ccaggatcac cacatagaat aggaaccgtg aggtacatgt ggatgtgcag 120

catggcactc acaaatttgc agaagggcag cccaaacatc caagtcttct tgatgaggta 180

ggtcaagcga aatggcactg tcagcagaaa aacgctgtgg accaccacca agttaatgac 240

cgccatggtg gtcactgacc gggtgttcat tttcaccagg aggaaaagaa tggaaatgac 300

acccaccagc ccgccaataa gcactatgaa gtagaggctg attaagtggg gtgtcactat 360

acgatcgcaa gaggaattcc tggaggtatt gtggccaggc atacttggga agtcacctgg 420

acgagaaaaa gcaccagagt aactgac 447

<210> 22
<211> 149
<212> PRT
<213> H.Sapiens

<400> 22

Val Ser Tyr Ser Gly Ala Phe Ser Pro Pro Gly Asp Phe Pro Ser Met
1 5 10 15

Pro Gly His Asn Thr Ser Arg Asn Ser Ser Cys Asp Pro Ile Val Thr
20 25 30

Pro His Leu Ile Ser Leu Tyr Phe Ile Val Leu Ile Gly Gly Leu Val
35 40 45

Gly Val Ile Ser Ile Leu Phe Leu Leu Val Lys Met Asn Thr Arg Ser
50 55 60

Val Thr Thr Met Ala Val Ile Asn Leu Val Val Val His Ser Val Phe
65 70 75 80

Leu Leu Thr Val Pro Phe Arg Leu Thr Tyr Leu Ile Lys Lys Thr Trp
85 90 95

Met Phe Gly Leu Pro Phe Cys Lys Phe Val Ser Ala Met Leu His Ile
100 105 110

His Met Tyr Leu Thr Val Pro Ile Leu Cys Gly Asp Pro Gly His Gln
115 120 125

Ile Pro His Leu Leu Gln Val Gln Arg Gln Ser Gly Ile Leu Gln Lys
130 135 140

Thr Ala Cys Cys Gly
145

<210> 23
<211> 222

<212> DNA
<213> H.Sapiens

<400> 23
actgaccaag gtcagggcat cgactgaggc tagaaggcca caggaaatgc cagtcaaggt 60

gttggcgcct gcaatcgcac ctaccacaaa cttgaccggg ggcagggggg caggcccgcc 120

agcgaacacg gtcagcagca ccagtccatt gcagagcacg gagagcaaca cgatggccca 180

cacagccagg cggatgcccc agctttcaaa gaggtactca ca 222

<210> 24
<211> 74
<212> PRT
<213> H.Sapiens

<400> 24

Cys Glu Tyr Leu Phe Glu Ser Trp Gly Ile Arg Leu Ala Val Trp Ala
1 5 10 15

Ile Val Leu Leu Ser Val Leu Cys Asn Gly Leu Val Leu Leu Thr Val
20 25 30

Phe Ala Gly Gly Pro Ala Pro Leu Pro Pro Val Lys Phe Val Val Gly
35 40 45

Ala Ile Ala Gly Ala Asn Thr Leu Thr Gly Ile Ser Cys Gly Leu Leu
50 55 60

Ala Ser Val Asp Ala Leu Thr Leu Val Ser
65 70

<210> 25
<211> 246
<212> DNA
<213> H.Sapiens

<400> 25
gaccgcatca tctacacgct caccaaccgc gacctgcgcc acgcgctcct gcgcctggtc 60

tgctgcgaac gccactcctg cggcagagac ccgagtggct cccagcagtc ggcgagcgcg 120

gctgaggctt ccgggggcct gcgccgctgc ctgcccccgg gccttgatgg gagcttcagc 180

ggctcggagc gctcatcgcc ccagcgcgac gggctggaca ccagcggctc cacaggcagc 240

cgcgct 246

<210> 26
<211> 82
<212> PRT
<213> H.Sapiens

<400> 26

Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg His Ala Leu
1 5 10 15

Leu Arg Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg Asp Pro Ser
20 25 30

Gly Ser Gln Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly Gly Leu Arg
35 40 45

Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly Ser Glu Arg
50 55 60

Ser Ser Pro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser Thr Gly Ser
65 70 75 80

Pro Gly

<210> 27
<211> 420
<212> DNA
<213> H.Sapiens

<220>
<221> misc_feature
<222> (81)..(106)
<223> n is any nucleic acid

<400> 27
cgtgaagaac agcgccacca tgaccagcat gtgcaccacg cgcgctctgc gccgcgatgc 60

tcgcgggtcc gcagcctcct nnnnnnnnnn nnnnnnnnnn nnnnnntggc agagcttgcg 120

cgcgatgcgg gcgtacatga ccacgatgag cgccagcggc gccaggtaga tgtgcgagaa 180

aagcacagtg gtgtagaccc tgcgcatgcc cttctcgggc caggcctccc agcaggagta 240

aagaaggtag gagcggttgc gggcgtccac catgaagtgg tgctcctcac gggtgacggt 300

cagcgtgacg gccgagggac acatgatgag cagcgccagg gcccagatga cggcgatggt 360

aacgagcgcc ttccgcaggg tcagcttctc gcggaaaggg tgcacgatgc agcggaacct 420

<210> 28
<211> 139
<212> PRT
<213> H.Sapiens

<220>
<221> UNSURE
<222> (104)..(113)
<223> Xaa is Unknown

<400> 28

Phe Arg Cys Ile Val His Pro Phe Arg Glu Lys Leu Thr Leu Arg Lys

1 5 10 15

Ala Leu Val Thr Ile Ala Val Ile Trp Ala Leu Ala Leu Leu Ile Met
20 25 30

Cys Pro Ser Ala Val Thr Leu Thr Val Thr Arg Glu Glu His His Phe
35 40 45

Met Val Asp Ala Arg Asn Arg Ser Tyr Pro Leu Tyr Ser Cys Trp Glu
50 55 60

Ala Trp Pro Glu Lys Gly Met Arg Arg Val Tyr Thr Thr Val Leu Phe
65 70 75 80

Ser His Ile Tyr Leu Ala Pro Leu Ala Leu Ile Val Val Met Tyr Ala
85 90 95

Arg Ile Ala Arg Lys Leu Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Glu Ala Ala Asp Pro Arg Ala Ser Arg Arg Arg Ala Arg Val Val
115 120 125

His Met Leu Val Met Val Ala Leu Phe Phe Thr
130 135

<210> 29
<211> 318
<212> DNA
<213> H.Sapiens

<400> 29
gcagcgggcg tgagtcctca ggcacttctt gaggtccttg ttgagcagga agcagacaat 60

tgggttgacg gcagcctggg cgaagctcat ccaaacagca gtggccaggt agcggtgggg 120

cacaccacag gctttcacaa acactcgcca gtagcaggcc acgatgtagg gtgaccagag 180

gagcagaaag agcagtgtga tcgcgtagaa catgcggccc agctgctttt caccсttgac 240

ctcgtccatg cccagtagcc gccggctggc tgcatgccca ttctgccgga tacccagcag 300

ggttggtggc atgggccc 318

<210> 30
<211> 106
<212> PRT
<213> H Sapiens

<400> 30

Gly Pro Met Pro Pro Thr Leu Leu Gly Ile Arg Gln Asn Gly His Ala
1 5 10 15

Ala Ser Arg Arg Leu Leu Gly Met Asp Glu Val Lys Gly Glu Lys Gln
20 25 30

Leu Gly Arg Met Phe Tyr Ala Ile Thr Leu Leu Phe Leu Leu Leu Trp

```
        35                  40                  45

Ser Pro Tyr Ile Val Ala Cys Tyr Trp Arg Val Phe Val Lys Ala Cys
    50                  55                  60

Ala Val Pro His Arg Tyr Leu Ala Thr Ala Val Trp Met Ser Phe Ala
65                  70                  75                  80

Gln Ala Ala Val Asn Pro Ile Val Cys Phe Leu Leu Asn Lys Asp Leu
                85                  90                  95

Lys Lys Cys Leu Arg Thr His Ala Pro Cys
            100                 105

<210>  31
<211>  354
<212>  DNA
<213>  H.Sapiens

<400>  31
tattctgtaa tgaagaatgt cattcacact gccattggca catccagtgg cctcacctag      60

cattgtgaaa gcccttcggt tggtgtattg ccacttcatt ttaaaaggat gcacaagtcc     120

ctggtgcctt tccacagcaa tgcaggtcat agtgaggatt tctgtcacaa cagcggtaga     180

ctggacaaat ggcaccatct tgcaaatgaa accacctgca gtaaggaaat aggataaatc     240

atacatcaaa acaaaaagaa taaaggtttc atctgtgtct ttgtaattat cactatcagt     300

ccattctgag cctctgccaa aaagtttgat aattgtaatt actctgtaga caca           354


<210>  32
<211>  117
<212>  PRT
<213>  H.Sapiens

<400>  32

Val Tyr Arg Val Ile Thr Ile Ile Lys Leu Phe Gly Arg Gly Ser Glu
1               5                   10                  15

Trp Thr Asp Ser Asp Asn Tyr Lys Asp Thr Asp Glu Thr Phe Ile Leu
            20                  25                  30

Phe Val Leu Met Tyr Asp Leu Ser Tyr Phe Leu Thr Ala Gly Ala Phe
        35                  40                  45

Ile Cys Lys Met Val Pro Phe Val Gln Ser Thr Ala Val Val Thr Glu
    50                  55                  60

Ile Leu Thr Met Thr Cys Ile Ala Val Glu Arg His Gln Gly Leu Val
65                  70                  75                  80

His Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg Arg Ala Phe Thr
                85                  90                  95

Met Leu Gly Glu Ala Thr Gly Cys Ala Asn Gly Ser Val Asn Asp Ile
```

100 105 110

Leu His Tyr Arg Ile
115

<210> 33
<211> 621
<212> DNA
<213> H.Sapiens

<400> 33
gagcaacatg atctttttga agtacttgac ggtgtcgttc ttgacggtca cgaagcacag 60

agtgttgatc atgctgttgc tcatggcgat gcactcgacg atgtagaagg cagtgaggta 120

gtgcttctcc ttcacaaaca cggtggggaa gaagtcgcgc acgatggtga agccgtagaa 180

gagcgcccag catagcacgt aggcggtgag gatgcacatg agcaccagga ccgtcttcct 240

gcggcagcgc agcctcttgc ggatctgctc tgtctggaat ccagggaccg ccttgaacca 300

gagctcccgg gagatcctgg catagcacag ggtcatggtg accacggggc ccacgaattc 360

tatgccaaag ataaagagga agtaggactt gtagtagagc tgctggtcca caggccagat 420

ctggccgcag aagatctttt cctggctctt gacaatgacg aggaccgtct cggtggtgaa 480

gtaggcggaa gggatggcga tcaggatgga caccgtccac accacggcaa tcaggccagt 540

ggctgtttgg cacttcattc gtggtctcag cggatggaca atagccagat acctagggca 600

agaacacaag tggaggcagc c 621

<210> 34
<211> 207
<212> PRT
<213> H.Sapiens

<400> 34

Gly Cys Leu His Leu Cys Ser Cys Pro Arg Tyr Leu Ala Ile Val His
1 5 10 15

Pro Leu Arg Pro Arg Met Lys Cys Gln Thr Ala Thr Gly Leu Ile Ala
20 25 30

Leu Val Trp Thr Val Ser Ile Leu Ile Ala Ile Pro Ser Ala Tyr Phe
35 40 45

Thr Thr Glu Thr Val Leu Val Ile Val Lys Ser Gln Glu Lys Ile Phe
50 55 60

Cys Gly Gln Ile Trp Pro Val Asp Gln Gln Leu Tyr Tyr Lys Ser Tyr
65 70 75 80

Phe Leu Phe Ile Phe Gly Ile Glu Phe Val Gly Pro Val Val Thr Met
85 90 95

Thr Leu Cys Tyr Ala Arg Ile Ser Arg Glu Leu Trp Phe Lys Ala Val
100 105 110

Pro Gly Phe Gln Thr Glu Gln Ile Arg Lys Arg Leu Arg Cys Arg Arg
115 120 125

Lys Thr Val Leu Val Leu Met Cys Ile Leu Thr Ala Tyr Val Leu Cys
130 135 140

Trp Ala Pro Phe Tyr Gly Phe Thr Ile Val Arg Asp Phe Phe Pro Thr
145 150 155 160

Val Phe Val Lys Glu Lys His Tyr Leu Thr Ala Phe Tyr Ile Val Glu
165 170 175

Cys Ile Ala Met Ser Asn Ser Met Ile Asn Thr Leu Cys Phe Val Thr
180 185 190

Val Lys Asn Asp Thr Val Lys Tyr Phe Lys Lys Ile Met Leu Leu
195 200 205

<210> 35
<211> 483
<212> DNA
<213> H.Sapiens

<400> 35
cagccacact gcagtgatga aatcaaatgt ccaacaccaa ccatagtcac cattactaac 60

taagaagcca caaaacttcc cttccagggt gttcagcagc agggacaggg cccagggcag 120

ggcacacatg acagttgaca ggtttcttgg gcagcagcag cagtaccaga taggccgcag 180

gacagacagg cagcactcag tactgatggc actcagcatg ctcaggccta caaggtaggc 240

aaaggtcatc acgctggtga agaagctagg gaaattgatg gagatggaac agaagaagtt 300

actgaggtac accaggcaat ttataatctg gaagcagagg aagaggaagt cggccccggc 360

caggctgagg acgtagacag agaaggcgtt cctgcgcatg cggaagccca ggagccagag 420

cacaaacccg tttcctacca gcccgaccag ggcaatgaaa aggatcagga agaccgggat 480

cag 483

<210> 36
<211> 161
<212> PRT
<213> H.Sapiens

<400> 36

Leu Ile Pro Val Phe Leu Ile Leu Phe Ile Ala Leu Val Gly Leu Val
1 5 10 15

Gly Asn Gly Phe Val Leu Trp Leu Leu Gly Phe Arg Met Arg Arg Asn
20 25 30

Ala Phe Ser Val Tyr Val Leu Ser Leu Ala Gly Ala Asp Phe Leu Phe
35 40 45

Leu Cys Phe Gln Ile Ile Asn Cys Leu Val Tyr Leu Ser Asn Phe Phe
50 55 60

Cys Ser Ile Ser Ile Asn Phe Pro Ser Phe Phe Thr Ser Val Met Thr
65 70 75 80

Phe Ala Tyr Leu Val Gly Leu Ser Met Leu Ser Ala Ile Ser Thr Glu
85 90 95

Cys Cys Leu Ser Val Leu Arg Pro Ile Trp Tyr Cys Cys Cys Cys Pro
100 105 110

Arg Asn Leu Ser Thr Val Met Cys Ala Leu Pro Trp Ala Leu Ser Leu
115 120 125

Leu Leu Asn Thr Leu Glu Gly Lys Phe Cys Gly Phe Leu Val Ser Asn
130 135 140

Gly Asp Tyr Gly Trp Cys Trp Thr Phe Asp Phe Ile Thr Ala Val Trp
145 150 155 160

Leu

<210> 37
<211> 330
<212> DNA
<213> H.Sapiens

<400> 37
gagagtctga ttctgactta catcacatat gtaggcctgg gcatttctat ttgcagcctg 60

atcctttgct tgtccgttga ggtcctagtc tggagccaag tgacaaagac agagatcacc 120

tatttacgcc atgtgtgcat tgttaacatt gcagccactt tgctgatggc agatgtgtgg 180

ttcattgtgg cttcctttct tagtggccca ataacacacc acaagggatg tgtggcagcc 240

acattttttg gtcatttctt ttacctttct gtatttttct ggatgcttgc caaggcactc 300

cttatcctct atggaatcat gattgttttc 330

<210> 38
<211> 110
<212> PRT
<213> H.Sapiens

<400> 38

Glu Ser Leu Ile Leu Thr Tyr Ile Thr Tyr Val Gly Leu Gly Ile Ser
1 5 10 15

Ile Cys Ser Leu Ile Leu Cys Leu Ser Val Glu Val Leu Val Trp Ser
20 25 30

Gln Val Thr Lys Thr Glu Ile Thr Tyr Leu Arg His Val Cys Ile Val
35 40 45

Asn Ile Ala Ala Thr Leu Leu Met Ala Asp Val Trp Phe Ile Val Ala
50 55 60

Ser Phe Leu Ser Gly Pro Ile Thr His His Lys Gly Cys Val Ala Ala
65 70 75 80

Thr Phe Phe Gly His Phe Phe Tyr Leu Ser Val Phe Phe Trp Met Leu
85 90 95

Ala Lys Ala Leu Leu Ile Leu Tyr Gly Ile Met Ile Val Phe
100 105 110

<210> 39
<211> 628
<212> DNA
<213> H.Sapiens

<400> 39
ttgtgtggca gtagagagat gtcaggcttc agagtcaaca agaactggat ttcaaactgg 60

atttgaggac ccccaccttt ggtaagtgac ttattatctg cgagcctctg tttctctctt 120

ctttaaatga ggacagtaaa tcccatacgg cagggtggtg gggagaatca gagatgatac 180

agctggtgat cacatctggt ttgtgttccc aggggcacca gactagggtt tctgagcatg 240

gatccaaccg tcccagtctt cggtacaaaa ctgacaccaa tcaacggacg tgaggagact 300

ccttcctaca atcagaccct gagcttcacg gtgctgacgt gcatcatttc ccttgtcgga 360

ctgacaggaa acgcggtagt gctctggctc ctgggctacc gcatgcgcag gaacgctgtc 420

tccatctaca tcctcaacct ggccgcagca gacttcctct tcctcagctt ccagattata 480

cgttcgccat tacgcctcat caatatcagc catctcatcc gcaaaatcct cgtttctgtg 540

atgacctttc cctactttac aggcctgagt atgctgagcg ccatcagcac cgagcgctgc 600

ctgtctgttc tgtggcccat ctggtacc 628

<210> 40
<211> 205
<212> PRT
<213> H.Sapiens

<400> 40

Leu Cys Gly Ser Arg Glu Met Ser Gly Phe Arg Val Asn Lys Asn Trp
1 5 10 15

Ile Ser Asn Trp Ile Gly Pro Pro Pro Leu Val Ser Asp Leu Leu Ser
20 25 30

Ala Ser Leu Cys Phe Ser Leu Leu Met Arg Thr Val Asn Pro Ile Arg
35 40 45

Gln Gly Gly Gly Glu Asn Gln Arg Tyr Ser Trp Ser His Leu Val Cys
50 55 60

Val Pro Arg Gly Thr Arg Leu Gly Phe Leu Ser Met Asp Pro Thr Val
65 70 75 80

Pro Val Phe Gly Thr Lys Leu Thr Pro Ile Asn Gly Arg Glu Glu Thr
85 90 95

Pro Cys Tyr Asn Gln Thr Leu Ser Phe Thr Val Leu Thr Cys Ile Ile
100 105 110

Ser Leu Val Gly Leu Thr Gly Asn Ala Val Val Leu Trp Leu Leu Gly
115 120 125

Tyr Arg Met Arg Arg Asn Ala Val Ser Ile Tyr Ile Leu Asn Leu Ala
130 135 140

Ala Ala Asp Phe Leu Phe Leu Ser Phe Gln Ile Ile Arg Ser Pro Leu
145 150 155 160

Arg Leu Ile Asn Ile Ser His Leu Ile Arg Lys Ile Leu Val Ser Val
165 170 175

Met Thr Phe Pro Tyr Phe Thr Gly Leu Ser Met Leu Ser Ala Ile Ser
180 185 190

Thr Glu Arg Cys Leu Ser Val Leu Trp Pro Ile Trp Tyr
195 200 205

<210> 41
<211> 319
<212> DNA
<213> H.Sapiens

<400> 41
acagaaagca aggccaccag gaccttaggc atagtcatgg gagtgtttgt gttgtgctgg 60

ctgcccttct ttgtcttgac gatcacagat cctttcatta attttacaac ccttgaagat 120

ctgtacaatg tcttcctctg gctaggctat ttcaactctg ctttcaatcc cattttatat 180

ggcatgcttt atccttggtt tcgcaaggca ttgaggatga ttgtcacagg catgatcttc 240

aaccctgact cttccaccct aagcctgttt tctgcccatg cttaggctgt gttcatcatt 300

caataggact cttttctgg 319

<210> 42
<211> 103
<212> PRT
<213> H.Sapiens

<400> 42

Thr Glu Ser Lys Ala Thr Arg Thr Leu Gly Ile Val Met Gly Val Phe
1 5 10 15

Val Leu Cys Trp Leu Pro Phe Phe Val Leu Thr Ile Thr Asp Pro Phe
20 25 30

Ile Asn Phe Thr Thr Leu Glu Asp Leu Tyr Asn Val Phe Leu Trp Leu
35 40 45

Gly Tyr Phe Asn Ser Ala Phe Asn Pro Ile Leu Tyr Gly Met Leu Tyr
50 55 60

Pro Trp Phe Arg Lys Ala Leu Arg Met Ile Val Thr Gly Met Ile Phe
65 70 75 80

His Pro Asp Ser Ser Thr Leu Ser Leu Phe Ser Ala His Ala Ala Val
85 90 95

Phe Ile Ile Gln Asp Ser Phe
100

<210> 43
<211> 515
<212> DNA
<213> H.Sapiens

<400> 43
taggaatctc agagaagaaa gtaaggaacc agaaaaccat aaaagaatgt aaatggaaaa 60

gaatcagcaa atcttattca cttatcacta aatctaaaat atgtcaaaat acatgaagac 120

aacaaatgct ttagaacaac tgttgaatgt attgtcctac aacttggcat atgatcatgc 180

ttgcctctct atgtccaagt gtttattttt gcagttgacc ttaatttcaa gttagttttg 240

aggtctctac agtaatgttt ttaatctgtc tctacttctt cagaaaataa attagttgtt 300

gacgaatcag tccttaagac cttgccgctt acaataagtt ttattgcctt cccaaaccat 360

tcgtaaaaga aagcataaat caaggggttc atagctgaat tataataaac acaccaaact 420

aaaatctcat aaacataagg aggagttata aaattcatat aagcatcaat cactgcatca 480

acgaggtatg gtagccaaga gacaagaaat gctgc 515

<210> 44
<211> 148
<212> PRT
<213> H.Sapiens

<400> 44

Leu His Gln Arg Gly Met Val Ala Lys Arg Gln Glu Met Leu Ala Ala
1 5 10 15

Phe Leu Val Ser Trp Leu Pro Tyr Leu Val Asp Ala Val Ile Asp Ala
20 25 30

Tyr Met Asn Phe Ile Thr Pro Pro Tyr Val Tyr Glu Ile Leu Val Trp
35 40 45

Cys Val Tyr Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr Ala Phe Phe
    50   55   60

Tyr Gln Trp Phe Gly Lys Ala Ile Lys Leu Ile Val Ser Gly Lys Val
65   70   75   80

Leu Arg Thr Asp Ser Ser Thr Thr Asn Leu Phe Ser Glu Glu Val Glu
                85   90   95

Thr Asp Lys His Tyr Cys Arg Asp Leu Lys Thr Asn Leu Lys Leu Arg
            100   105   110

Ser Thr Ala Lys Ile Asn Thr Trp Thr Arg Gly Lys His Asp His Met
        115   120   125

Pro Ser Cys Arg Thr Ile His Ser Thr Val Val Leu Lys His Leu Leu
    130   135   140

Ser Ser Cys Ile
145

<210> 45
<211> 726
<212> DNA
<213> H.Sapiens

<400> 45
ctggaaagag gtcctcgatc tatcctctac gccgtccttg gttttggggc tgtgctggca 60

gtgtttggaa acttactggt catgattgct atccttcact tctaacaact gcacacacct 120

acaaactttc tgattgcgtc gctggcctgt gctgacttct tggtgggagt cactgtgatg 180

cacttcagca cagtgaggtc tgtggagagc tgttggtact ttggggacag ttactgtaaa 240

ttccatacat gttttgacac atctttctgt tttgcttctt tatttcattt atgctgtatc 300

tttgttgata gatacattgc tgttactgat cctctgacct atccaaccaa gtttactgtg 360

ttagtttcag ggatatgcat tgttctttcc tggttctttt ctgtcacata cagcttttcg 420

atcttttaca cgggagccaa cgaagaagga attgaggaat tagtagttgc tctaacctgt 480

gtaggaggct gccaggctcc actgaatcaa aactgggtcc tactttgttt tcttctattc 540

tttataccca atgtcgccat ggtgtttata tacagtaaga tatttttggt ggccaagcat 600

caggctagga agatagaaag tacagccagc caagctcagt ccttctcaga gagttacaag 660

gaaagagtag caaaaagaga gagaaaggct gccaaaacct tgggaattgc tatggcagca 720

cttctt 726

<210> 46
<211> 241
<212> PRT
<213> H.Sapiens

<400> 46

Leu Glu Arg Gly Pro Arg Ser Ile Leu Tyr Ala Val Leu Gly Phe Gly
1   5   10   15

Ala Val Leu Ala Ala Phe Gly Asn Leu Leu Val Met Ile Ala Ile Leu
            20   25   30

His Phe Gln Leu His Thr Pro Thr Asn Phe Leu Ile Ala Ser Leu Ala
        35   40   45

Cys Ala Asp Phe Leu Val Gly Val Thr Val Met Pro Phe Ser Thr Val
    50   55   60

Arg Ser Val Glu Ser Cys Trp Tyr Phe Gly Asp Ser Tyr Cys Lys Phe
65   70   75   80

His Thr Cys Phe Asp Thr Ser Phe Cys Phe Ala Ser Leu Phe His Leu
                85   90   95

Cys Cys Ile Ser Val Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Thr
            100   105   110

Tyr Pro Thr Lys Phe Thr Val Ser Val Ser Gly Ile Cys Ile Val Leu
        115   120   125

Ser Trp Phe Phe Ser Val Thr Tyr Ser Phe Ser Ile Phe Tyr Thr Gly
    130   135   140

Ala Asn Glu Glu Gly Ile Glu Glu Leu Val Val Ala Leu Thr Cys Val
145   150   155   160

Gly Gly Cys Gln Ala Pro Leu Asn Gln Asn Trp Val Leu Leu Cys Phe
                165   170   175

Leu Leu Phe Phe Ile Pro Asn Val Ala Met Val Phe Ile Tyr Ser Lys
            180   185   190

Ile Phe Leu Val Ala Lys His Gln Ala Arg Lys Ile Glu Ser Thr Ala
        195   200   205

Ser Gln Ala Gln Ser Phe Ser Glu Ser Tyr Lys Glu Arg Val Ala Lys
    210   215   220

Arg Glu Arg Lys Ala Ala Lys Thr Leu Gly Ile Ala Met Ala Ala Phe
225   230   235   240

Leu

<210> 47
<211> 660
<212> DNA
<213> H.Sapiens

<400> 47
aaccaggtgg ccttactcct aagacccctg gccttgtcta tggcctttat caacagctgt 60

ctcaatccag ttctctatgt cttcattggg catgacttct gggagcactt gctccactcc 120

ctgctagctg ccttagaacg ggcacttagc gaggagccag atagtgcctg aatcccagct 180

cccaggcaga tgagtccttt ataacatgac ccaatttcct actccatttt cccaccactc 240

aatcctcttc ccaaacagct ctaccataat ccaacatcca acagaattta agagaataaa 300

ccacaacttt taagtgagct ctatgtgcta ggtcatgttt tagaatacaa ccttaagtgc 360

ctggaagatg gaggcaagaa acaaacaagg tctcattctt tagaggaaga cagttcacca 420

agactcaaac agaaaaaaag atagttatct tgtgacaaaa caagtcataa aattgggtca 480

ggacctgcag caatgacttt atgctagaat ccagagcact agcaggaaac tgcttaaatt 540

ttacttaatc aaagtcaagt ttggacatac atgtcaggta aaacctagca gagatgagct 600

accttgattt taaaacttca agggatagct caatgtcatc aagatccttt tgatgacttg 660

<210> 48
<211> 211
<212> PRT
<213> H.Sapiens

<400> 48

Asn Gln Val Ala Leu Leu Leu Arg Pro Leu Ala Leu Ser Met Ala Phe
1 5 10 15

Ile Asn Ser Cys Leu Asn Pro Val Leu Tyr Val Phe Ile Gly His Asp
20 25 30

Phe Trp Glu His Leu Leu His Ser Leu Leu Ala Ala Leu Glu Arg Ala
35 40 45

Leu Ser Glu Glu Pro Asp Ser Ala Ile Pro Ala Pro Arg Gln Met Ser
50 55 60

Pro Leu His Asp Pro Ile Ser Tyr Ser Ile Phe Pro Pro Leu Asn Pro
65 70 75 80

Leu Pro Lys Gln Leu Tyr His Asn Pro Thr Ser Asn Arg Ile Glu Asn
85 90 95

Lys Pro Gln Leu Leu Ser Glu Leu Tyr Val Leu Gly His Val Leu Glu
100 105 110

Tyr Asn Leu Lys Cys Leu Glu Asp Gly Gly Lys Lys Gln Thr Arg Ser
115 120 125

His Ser Leu Glu Glu Asp Ser Ser Pro Arg Leu Lys Gln Lys Lys Arg
130 135 140

Leu Ser Cys Asp Lys Thr Ser His Lys Ile Gly Ser Gly Pro Ala Ala
145 150 155 160

Met Thr Leu Cys Asn Pro Glu His Gln Glu Thr Ala Ile Leu Leu Asn

```
                165                 170                 175

Gln Ser Gln Val Trp Thr Tyr Met Ser Gly Lys Thr Gln Arg Ala Thr
            180                 185                 190

Leu Ile Leu Lys Leu Gln Gly Ile Ala Gln Cys His Gln Asp Pro Phe
        195                 200                 205

Asp Asp Leu
    210

<210>  49
<211>  465
<212>  DNA
<213>  H.Sapiens

<400>  49
gcttgttcac ggccaccatc ctcaagctgt tgcgcacgga ggaggcgcac ggccgggagc      60

agcggaggcg cgcggtgggc ctggccgcgg tggtcttgct ggcctttgtc acctgcttcg     120

cccccaacaa cttcgtgctc ctggcgcaca tcgtgagccg cctgttctac ggcaagagct     180

actaccacgt gtacaagctc acgctgtgtc tcagctgcct caacaactgt ctggacccgt     240

ttgtttatta ctttgcgtcc cgggaattcc agctgcgcct gcgggaatat ttgggctgcc     300

gccgcgtgcc cagagacacc ctggacacgc gccgcgagag cctcttctcc gccaggacca     360

catccgtgcg ctccgaggcc ggtgcgcacc ctgaagggat ggagggagcc accaggcccg     420

gcctccagag gcaggagagt gtgttctgag tcccgggggc gcagc                     465


<210>  50
<211>  160
<212>  PRT
<213>  H.Sapiens

<400>  50

Leu Phe Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu Glu Ala His
1               5                   10                  15

Gly Arg Glu Gln Arg Arg Arg Ala Val Gly Leu Ala Ala Val Val Leu
            20                  25                  30

Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val Leu Leu Ala
        35                  40                  45

His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr His Val Tyr
    50                  55                  60

Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu Asp Pro Phe
65                  70                  75                  80

Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu Arg Glu Tyr
                85                  90                  95
```

Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr Arg Arg Glu
100 105 110

Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu Ala Gly Ala
115 120 125

His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu Gln Arg Gln
130 135 140

Glu Ser Val Phe Val Pro Gly Ala Gln Ala Ala Pro Pro Gly Leu Arg
145 150 155 160

<210> 51
<211> 603
<212> DNA
<213> H.Sapiens

<400> 51
ttacttattc tgccctttat ccaactttta attccctttg ctattctcct gcctcatttt 60

ctggcctcat tttccctatt atcctgcctc acattgatca agggatgagg ctggcaggat 120

ccggaaccca cagggccccg tgggccatga gaggctcctg gacttgaacc tcaggacact 180

ccaactctgg ctgccggcag ggatggaagc tggatgagca ggcaggagct ggcagtgggg 240

gtggagagcc ataggctatt gggatggaca ggcttgggtg cctcatggga gctccccatg 300

ggagctgtgg ccccttgggg cctcctattt ctcaccccag gctttcccgg gagaggctca 360

agtcagaaga tgccccaaag atccacgtgg ccctgggtgg cagcctgttc ctcctgaatc 420

tggccttctt ggtcaatgtg gggagtggct caaaggggtc tgatgctgcc tgctgggccc 480

ggggagctct cttccactac ttcctgctct gtgccttcac ctggatgggc cttgaagcct 540

tccacctcta cctgctcgct gtcagggtct tcaacaccta cttcgggcac tacttcctga 600

agc 603

<210> 52
<211> 198
<212> PRT
<213> H.Sapiens

<400> 52

Glu Thr Tyr Ser Ala Leu Tyr Pro Thr Phe Asn Ser Leu Cys Tyr Ser
1 5 10 15

Pro Ala Ser Phe Ser Gly Leu Ile Phe Pro Ile Ile Leu Pro His Ile
20 25 30

Asp Gln Gly Met Arg Leu Ala Gly Ser Gly Thr His Arg Ala Pro Trp
35 40 45

Ala Met Arg Gly Ser Trp Thr Thr Ser Gly His Ser His Ser Gly Cys
50 55 60

Arg Gln Gly Trp Lys Leu Asp Glu Gln Ala Gly Ala Gly Ser Gly Gly
65 70 75 80

Gly Glu Pro Ala Ile Gly Val Asp Arg Leu Gly Cys Leu Met Gly Ala
85 90 95

Pro His Gly Ser Cys Gly Pro Leu Gly Pro Leu Ile Ser His Pro Arg
100 105 110

Leu Ser Arg Glu Arg Phe Lys Ser Glu Asp Ala Pro Lys Ile His Val
115 120 125

Ala Leu Gly Gly Ser Leu Phe Leu Leu Asn Leu Ala Phe Leu Val Asn
130 135 140

Val Gly Ser Gly Ser Lys Gly Ser Asp Ala Ala Cys Trp Ala Arg Gly
145 150 155 160

Ala Val Phe His Tyr Phe Leu Leu Cys Ala Phe Thr Trp Met Gly Leu
165 170 175

Glu Ala Phe His Leu Tyr Leu Leu Ala Val Arg Val Phe Asn Thr Tyr
180 185 190

Phe Gly His Tyr Phe Leu
195

<210> 53
<211> 335
<212> DNA
<213> H.Sapiens

<400> 53
aattggtcgg agagtgcagc tgcttgaaat ggaggattga aatcatcacc aggaggtttc 60

caaacacagc cagcacagcc ccaaagccaa acactatgta cagaatcacc cgggatcccg 120

gcgagaaggg gattttcaca caggacccat tcacgttcgc gtagcacagc tgcacagcca 180

ccagcaggga tgaattgctg ctcataacgc tggtatttac atatggagaa attttgtcct 240

tgttcattat cacaaaaaat acaggattgt tcctgatttt cattgctcct gccgaaaaaa 300

acacatattc accaggatgc cagaggaaat gatca 335

<210> 54
<211> 111
<212> PRT
<213> H.Sapiens

<400> 54

Asp His Phe Leu Trp His Pro Gly Glu Tyr Val Phe Phe Ser Ala Gly
1 5 10 15

Ala Met Lys Ile Arg Asn Asn Pro Val Phe Phe Val Ile Ile Asn Lys
20 25 30

Asp Lys Ile Ser Pro Tyr Val Asn Thr Ser Val Met Ser Ser Asn Ser
35 40 45

Ser Leu Leu Val Ala Val Gln Leu Cys Tyr Ala Asn Val Asn Gly Ser
50 55 60

Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg Val Ile Leu Tyr Ile
65 70 75 80

Val Phe Gly Phe Gly Ala Val Leu Ala Val Phe Gly Asn Leu Leu Val
85 90 95

Met Ile Ser Ile Leu His Phe Lys Gln Leu His Ser Pro Thr Asn
100 105 110

<210> 55
<211> 586
<212> DNA
<213> H.Sapiens

<400> 55
cacatcttaa caagactgaa aaacattgat ttgtttttaa tttgaagagc aatttatttg 60

ctattcattc atagtcttac ttgattttta aaaactcatt tcgcttggta attttaaagg 120

tatcctgaac ttcgtctatc caactgctta tatatgttca gaaaacaaat tcatggttgc 180

tgaactgttc tttaaaacct gaccagttac aataactttt attgctttcc taaaccatgg 240

gtaagataaa gcataaatca aaggattcat ggctgagtta taataagcac accaacagca 300

tcatgaatac aggcaggggt tataaagccc ataaaggcat caattaatga atcaatgcta 360

tatgttaacc atgaaatcat aaatgctacc actgtgaccc ccaggttttt agctgctttt 420

ctctctctcc tggccactct ggctttgtaa ctctctgagg atgattctgt cttgctacca 480

gtattttcta tctttttcgc ctgtcgtcta gccacaagaa atatgttacc atacagaatt 540

atcataataa aggtaggtat aaagaaggat agaaaatctg tcaaca 586

<210> 56
<211> 190
<212> PRT
<213> H.Sapiens

<400> 56

Leu Thr Asp Phe Leu Ser Phe Phe Ile Pro Thr Phe Ile Met Ile Ile
1 5 10 15

Leu Tyr Gly Asn Ile Phe Leu Val Ala Arg Arg Gln Ala Lys Lys Ile
20 25 30

Glu Asn Thr Gly Ser Lys Thr Glu Ser Ser Ser Glu Ser Tyr Lys Ala
35 40 45

Arg Val Ala Arg Arg Glu Arg Lys Ala Ala Lys Thr Leu Gly Val Thr
50 55 60

Val Val Ala Phe Met Ile Ser Trp Leu Pro Tyr Ser Ile Asp Ser Leu
65 70 75 80

Ile Asp Ala Phe Met Gly Phe Ile Thr Pro Ala Cys Ile Tyr Glu Ile
85 90 95

Cys Cys Trp Cys Ala Tyr Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr
100 105 110

Ala Leu Phe Tyr Pro Trp Phe Arg Lys Ala Ile Lys Val Ile Val Thr
115 120 125

Gly Gln Val Leu Lys Asn Ser Ser Ala Thr Met Asn Leu Phe Ser Glu
130 135 140

His Ile Ala Val Gly Thr Lys Phe Arg Ile Pro Leu Lys Leu Pro Ser
145 150 155 160

Glu Met Ser Phe Lys Ser Ser Lys Thr Met Asn Glu Gln Ile Asn Cys
165 170 175

Ser Ser Asn Lys Gln Ile Asn Val Phe Gln Ser Cys Asp Val
180 185 190

<210> 57
<211> 976
<212> DNA
<213> H.Sapiens

<400> 57
tttgtggcaa ggagaccctg atcccggtct tcctgatcct tttcattgcc ctggtcgggc 60

tggtaggaaa cgggtttgtg ctctggctcc tgggcttccg catgcgcagg aacgccttct 120

ctgtctacgt cctcagcctg gccggggccg acttcctctt cctctgcttc cagattataa 180

attgcctggt gtacctcagt aacttcttct gttccatctc catcaatttc cctagcttct 240

tcaccactgt gatgacctgt gcctaccttg caggcctgag catgctgagc accgtcagca 300

ccgagcgctg cctgtccgtc ctgtggccca tctggtatcg ctgccgccgc cccagacacc 360

tgtcagcggt cgtgtgtgtc ctgctctggg ccctgtccct actgctgagc atcttggaag 420

ggaacttctg tggcttctta tttagtgatg gtgactctgg ttggtgtcag acatttgatt 480

tcatcactgc agcgtggctg atttttttat tcatggttct ctgtgggtcc agtctggccc 540

tgctggtcag gatcctctgt ggctccaggg gtctgccact gaccaggctg tacctgacca 600

tcctgctcac agtgctggtg tccctcctct gcggcctgcc ctttggcatt cagtggttcc 660

taatattatg gatctggaag gattctgatg tcttattttg tcatattcat ccagtttcag 720

ttgtcctgtc atctcttaac agcagtgcca accccatcat ttacttcttc gtgggctctt 780

ttaggaagca gtggcggstg cagcacccga tcctcaagct ggctctccag agggctctgc 840

aggacattgc tgaggtggat cacagtgaag gatgcttccg tcagggcacc cggagattca 900

aagaagcatt ctggtgtagg gatggacccc tctacttcca tcatatatat gtggctttga 960

gaggcaactt tgcccc 976

<210> 58
<211> 324
<212> PRT
<213> H.Sapiens

<220>
<221> UNSURE
<222> (266)..(266)
<223> Xaa is Unknown

<400> 58

Cys Gly Lys Glu Thr Leu Ile Pro Val Phe Leu Ile Leu Phe Ile Ala
1 5 10 15

Leu Val Gly Leu Val Gly Asn Gly Phe Val Leu Trp Leu Leu Gly Phe
20 25 30

Arg Met Arg Arg Asn Ala Phe Ser Val Tyr Val Leu Ser Leu Ala Gly
35 40 45

Ala Asp Phe Leu Phe Leu Cys Phe Gln Ile Ile Asn Cys Leu Val Tyr
50 55 60

Leu Ser Asn Phe Phe Cys Ser Ile Ser Ile Asn Phe Pro Ser Phe Phe
65 70 75 80

Thr Thr Val Met Thr Cys Ala Tyr Leu Ala Gly Leu Ser Met Leu Ser
85 90 95

Thr Val Ser Thr Glu Arg Cys Leu Ser Val Leu Trp Pro Ile Trp Tyr
100 105 110

Arg Cys Arg Arg Pro Arg His Leu Ser Ala Val Val Cys Val Leu Leu
115 120 125

Trp Ala Leu Ser Leu Leu Leu Ser Ile Leu Glu Gly Lys Phe Cys Gly
130 135 140

Phe Leu Phe Ser Asp Gly Asp Ser Gly Trp Cys Gln Thr Phe Asp Phe
145 150 155 160

Ile Thr Ala Ala Trp Leu Ile Phe Leu Phe Met Val Leu Cys Gly Ser
165 170 175

Ser Leu Ala Leu Leu Val Arg Ile Leu Cys Gly Ser Arg Gly Leu Pro
180 185 190

Leu Thr Arg Leu Tyr Leu Thr Ile Leu Leu Thr Val Leu Val Ser Leu

195 200 205

Leu Cys Gly Leu Pro Phe Gly Ile Gln Trp Phe Leu Ile Leu Trp Ile
210 215 220

Trp Lys Asp Ser Asp Val Leu Phe Cys His Ile His Pro Val Ser Val
225 230 235 240

Val Leu Ser Ser Leu Asn Ser Ser Ala Asn Pro Ile Ile Tyr Phe Phe
245 250 255

Val Gly Ser Phe Arg Lys Gln Trp Arg Xaa Gln His Pro Ile Leu Lys
260 265 270

Leu Ala Leu Gln Arg Ala Leu Gln Asp Ile Ala Glu Val Asp His Ser
275 280 285

Glu Gly Cys Phe Arg Gln Gly Thr Arg Arg Phe Lys Glu Ala Phe Trp
290 295 300

Cys Arg Asp Gly Pro Leu Tyr Phe His His Ile Tyr Val Ala Leu Arg
305 310 315 320

Gly Asn Phe Ala

<210> 59
<211> 578
<212> DNA
<213> H.Sapiens

<400> 59
ctttgcatct cactgttgag cagacagcct gctgaaagtt gtcgctgacc accacatata 60

gtaacaggtt atcaaaggtg ttcagagcag cataatggtc tagaaacgat gtaagcttca 120

tggatctgat tttcaatgga acaactgatt gaaagcaggc tgagattcga tcctgaatga 180

ccctcaagat atggaagggt aaaaaacata cgtaaaatgc aaggagtagc agaatggtta 240

gccttcgtgc tttctgctta aggcagctgt cagtttgcag tccatgggtc aaagtgtgga 300

taatcgtggt atagcaaagt gtcactatca ccaaggggag gcagaaagta cttgcagtca 360

aaatcaggtt gtaccactta atagtattga gttcatccga actggtgagg tcgagacagg 420

ctgatctgtt ggttctgttg gttgatgtga tcaagaaggt catcggaatg acagctacca 480

gtgaaatgat ccacaccaca gcacaggcta caactgcaca tcgagttttg tgaatggaaa 540

agcagctcat tgggtgaatg atcacacagt agcggaag 578

<210> 60
<211> 192
<212> PRT
<213> H.Sapiens

<400> 60

Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile His
1 5 10 15

Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile Ser
20 25 30

Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Thr Asn Arg
35 40 45

Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu Asn
50 55 60

Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Ser Thr Phe Cys Leu
65 70 75 80

Pro Leu Val Ile Val Thr Leu Cys Tyr Thr Thr Ile Ile His Thr Leu
85 90 95

Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ala Arg Arg
100 105 110

Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro Phe
115 120 125

His Ile Leu Arg Val Ile Gln Asp Arg Ile Ser Ala Cys Phe Gln Ser
130 135 140

Val Val Pro Leu Arg Ile Arg Ser Met Lys Leu Thr Ser Phe Leu Asp
145 150 155 160

His Tyr Ala Ala Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val Val
165 170 175

Val Ser Asp Asn Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys Lys
180 185 190

<210> 61
<211> 872
<212> DNA
<213> H.Sapiens

<400> 61
gggagggctc gtagacacac taaccctacc ctttctgttt cttcctcatc tttcctttcc 60

atctgtttct catggtctcc tgtctgtctc tctctctctc ccctctttct ctctcctcgc 120

tctttctcat cccctccatt tctgtgtcaa tctcaatcca tttatatcgg tggccacttt 180

tctatctctt tgttctatct ctctctctct ctctttccca ctttgtctct gcacgcctgt 240

tgtgtttttc tgcctgtctc tctcttgccc tcatctctct gtctctctct tgccctcatc 300

tctctgtctc tctgtgtctg tgtctccccc gctcattccc atttgcaggt gcaatgtagc 360

aggacaactc atggagcccc cccgggccca tcgagtaccg gactggctga ccccctaggg 420

ttggcagtag cccctgaccc tcagtatggc caacactacc ggagagcctg aggaggtgag 480

cggcgctctg tccccaccgt ccgcatcagc ttatgtgaag ctggtactgc tgggactgat 540

tatgtgcgtg agcctggcgg gtaacgccat cttgtccctg ctggtgctca aggagcgggc 600

cctgcacaag gctccttact acttcctgct ggacctgtgc ctggccgatg gcatacgctc 660

tgccgtctgc ttcccctttg tgctggcttc tgtgcgccac ggctcttcat ggaccttcag 720

tgcactcagc tgcaagattg tggcctttat ggccgtgctc ttttgcttcc atgcggcctt 780

catgctgttc tgcatcagcg tcacccgcta catggccatc gcccaccacc gcttctacgc 840

caagcgcatg acactctgga catgcgcggc tg 872

<210> 62
<211> 143
<212> PRT
<213> H.Sapiens

<400> 62

Met Ala Asn Thr Thr Gly Glu Pro Glu Glu Val Ser Gly Ala Leu Ser
1 5 10 15

Pro Pro Ser Ala Ser Ala Tyr Val Lys Leu Val Leu Leu Gly Leu Ile
20 25 30

Met Cys Val Ser Leu Ala Gly Asn Ala Ile Leu Ser Leu Leu Val Leu
35 40 45

Lys Glu Arg Ala Leu His Lys Ala Pro Tyr Tyr Phe Leu Leu Asp Leu
50 55 60

Cys Leu Ala Asp Gly Ile Arg Ser Ala Val Cys Phe Pro Phe Val Leu
65 70 75 80

Ala Ser Val Arg His Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys
85 90 95

Lys Ile Val Ala Phe Met Ala Val Leu Phe Cys Phe His Ala Ala Phe
100 105 110

Met Leu Phe Cys Ile Ser Val Thr Arg Tyr Met Ala Ile Ala His His
115 120 125

Arg Phe Tyr Ala Lys Arg Met Thr Leu Trp Thr Cys Ala Ala Glu
130 135 140

<210> 63
<211> 962
<212> DNA
<213> H.Sapiens

<400> 63
aaaattgct gtactgaact attgaatgga acttggaaat aaagtccctt ccaaaataac 60

tattcttcaa cagagagtaa taggtaaatg ttttagaagt gagaggactc aaattgccaa 120

tgatttactc ttttattttt cctcctaggt ttctgggata agtatgtgca aataaaaaat 180

aaacatgaga aggaactgta acctgattat ggatttggga aaaagataaa tcaacacaca 240

aagggaaaag taaactgatt gacagccctc aggaatgatg cccttttgcc acaatataat 300

taatatttcc tgtgtgaaaa acaactggtc aaatgatgtc cgtgcttccc tgtacagttt 360

aatggtgctc ataattctga ccacactcgt tggcaatctg atagttattg tttctatatc 420

acacttcaaa caacttcata ccccaacaaa ttggctcatt cattccatgg ccactgtgga 480

ctttcttctg ggtgtctgg tcatgcctta cagtatggtg agatctgctg agcactgttg 540

gtattttgga gaagtcttct gtaaaattca cacaagcacc gacattatgc tgagcttagc 600

ctccattttc catttgtctt tcatctccat tgaccgctac tatgctgtgt gtgatccact 660

gagatataaa gccaagatga atatcttggt tatttgtgtg atgatcttca ttagttggag 720

tgtccctgct gtttttgcat ttggaatgat ctttctggag ctaaacttca aaggcgctga 780

agagatatat tacaaacatg ttcactgcag aggaggttgc tctgtcttct ttagcaaaat 840

atctggggta ctgaccttta tgacttcttt ttatatacct ggatctatta tgttatgtgt 900

ctattacaga atatatctta tcgctaagga acaggcaaga ttaattagtg atgccaatca 960

ga 962

<210> 64
<211> 238
<212> PRT
<213> H.Sapiens

<400> 64

Arg Glu Lys Thr Asp Gln Pro Ser Gly Met Met Pro Phe Cys His Asn
1 5 10 15

Ile Ile Asn Ile Ser Cys Val Lys Asn Asn Trp Ser Asn Asp Val Arg
20 25 30

Ala Ser Leu Tyr Ser Leu Met Val Leu Ile Ile Leu Thr Thr Leu Val
35 40 45

Gly Asn Leu Ile Val Ile Val Ser Ile Ser His Phe Lys Gln Leu His
50 55 60

Thr Pro Thr Asn Trp Leu Ile His Ser Met Ala Thr Val Asp Phe Leu
65 70 75 80

Leu Gly Cys Leu Val Met Pro Tyr Ser Met Val Arg Ser Ala Glu His
85 90 95

Cys Trp Tyr Phe Gly Glu Val Phe Cys Lys Ile His Thr Ser Thr Asp
100 105 110

Ile Met Leu Ser Ser Ala Ser Ile Phe His Leu Ser Phe Ile Ser Ile
115 120 125

Asp Arg Tyr Tyr Ala Val Cys Asp Pro Leu Arg Tyr Lys Ala Lys Met
130 135 140

Asn Ile Leu Val Ile Cys Val Met Ile Phe Ile Ser Trp Ser Val Pro
145 150 155 160

Ala Val Phe Ala Phe Gly Met Ile Phe Leu Glu Leu Asn Phe Lys Gly
165 170 175

Ala Glu Glu Ile Tyr Tyr Lys His Val His Cys Arg Gly Gly Cys Ser
180 185 190

Val Phe Phe Ser Lys Ile Ser Gly Val Leu Thr Phe Met Thr Ser Phe
195 200 205

Tyr Ile Pro Gly Ser Ile Met Leu Cys Val Tyr Tyr Arg Ile Tyr Leu
210 215 220

Ile Ala Lys Glu Gln Ala Arg Leu Ile Ser Asp Ala Asn Gln
225 230 235

<210> 65
<211> 1018
<212> DNA
<213> H.Sapiens

<400> 65
aacaatcccg ggtggaacct gggcatgtat attttgattg ttttatgcat actcctagtg 60

aagaaccaat gtcttgctca gatagaagca agatactcag acttagtttc tctgtagctc 120

ctgcttttta ttattcctgg ttggattgca ccactactca gtttctattt tataatactg 180

attataaaac atgggaggga aataactttg tattggtttt tatggataat ttattatgtg 240

tcctagactc tggccttgtc aaaagaagga cgtaagaagg cacgatgtat tatacttggg 300

aatgatagaa gagactgacc tggtatttcc acccggaaga gggaaaggat tttaactaca 360

aatacaggaa tccagcagat ggcatcagag aacactataa aaaagaaacg atttgcaaca 420

gccactctct ttccaaaaca attccttact tctgtggtct gcaaggcggt tttttgaatg 480

gaacagaaca tagtaatata ggaaaacaca atgatgagaa aagccagcaa gttcacacct 540

gttggggaaa agcacacttt taacattcca ggcgtaaaag tcaacagtaa aattactgtg 600

gtacaggttg agtatccctt acccaaaatg tttgaaacca gaaatgtttt ggatttcgga 660

tttcggaata tttacacatt cataatgata tatcttggaa atggttccca agtctaaaca 720

taaaatttat ttatgtttca tatacacctt atacacatag tctgaaagta attttgtaca 780

atattttaaa taattttggg catgaaacaa agtttgcata cattgaacca tcagacagca 840

aaagcttcag gtgtggaatt ttccacttgt ggcatcatgt tgatgctcaa aaagttccat 900

attttagagc atttcaaatt ttggattttc aaattacaaa tgcttaacct gtacttagat 960

gttaaataca gtgcctcttc cacgggcact ttcaggaagc atttctttat ataagccc 1018

<210> 66
<211> 327
<212> PRT
<213> H.Sapiens

<400> 66

Tyr Ile Lys Glu Cys Phe Leu Lys Val Pro Val Glu Glu Ala Leu Tyr
1 5 10 15

Leu Thr Ser Lys Tyr Arg Leu Ser Ile Cys Asn Leu Lys Ile Gln Asn
20 25 30

Leu Lys Cys Ser Lys Ile Trp Asn Phe Leu Ser Ile Asn Met Met Pro
35 40 45

Gln Val Glu Asn Ser Thr Pro Glu Ala Phe Ala Val Trp Phe Asn Val
50 55 60

Cys Lys Leu Cys Phe Met Pro Lys Ile Ile Asn Ile Val Gln Asn Tyr
65 70 75 80

Phe Gln Thr Met Cys Ile Arg Cys Ile Asn Ile Asn Lys Phe Cys Val
85 90 95

Thr Trp Glu Pro Phe Pro Arg Tyr Ile Ile Met Asn Val Ile Phe Arg
100 105 110

Asn Pro Lys Ser Lys Thr Phe Leu Val Ser Asn Ile Leu Gly Lys Gly
115 120 125

Tyr Ser Thr Cys Thr Thr Val Ile Leu Leu Leu Thr Phe Thr Pro Glu
130 135 140

Met Leu Lys Val Cys Phe Ser Pro Thr Gly Val Asn Leu Leu Ala Phe
145 150 155 160

Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys
165 170 175

Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg Glu Val
180 185 190

Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala Ile Cys
195 200 205

Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu
210 215 220

Ile Pro Gly Gln Ser Leu Leu Ser Phe Pro Ser Ile Ile His Arg Ala
225 230 235 240

Phe Leu Arg Pro Ser Phe Asp Lys Ala Arg Val Asp Thr Ile Ile His

245 250 255

Lys Asn Gln Tyr Lys Val Ile Ser Leu Pro Cys Phe Ile Ile Ser Ile
260 265 270

Ile Lys Lys Leu Ser Ser Gly Ala Ile Gln Pro Gly Ile Ile Lys Ser
275 280 285

Arg Ser Tyr Arg Glu Thr Lys Ser Glu Tyr Leu Ala Ser Ile Ala Arg
290 295 300

His Trp Phe Phe Thr Arg Ser Met His Lys Thr Ile Lys Ile Tyr Met
305 310 315 320

Pro Arg Phe His Pro Gly Leu
325

<210> 67
<211> 1251
<212> DNA
<213> H.Sapiens

<400> 67
actaccatgg aagctgacct gggtgccact ggccacaggc cccgcacaga gcttgatgat 60

gaggactcct acccccaagg tggctgggac acggtcttcc tggtggccct gctgctcctt 120

gagctgccag ccaatgggtt gatggcgtgg ctggccggct cccaggcccg gcatggagct 180

gacacgcgtc tggcgctgct cctgctcagc ctggccctct ctgacttctt gttcctggca 240

gcagcggcct tccagatcct agagatccgg catggggac actggccgct ggggacagct 300

gcctgccgct tctactactt cctatggggc gtgtcctact cctccggcct cttcctgctg 360

gccgccctca gcctcgaccg ctgcctgctg gcgctgtgcc cacactggta ccctgggcac 420

cacccagtcc gcctgcccct ctgggtctgc gccggtgtct gggtgctggc cacactcttc 480

agcgtgccct ggctggtctt ccccgaggct gccgtctggt ggtacgacct ggtcatctgc 540

ctggacttct gggacagcga ggagctgtcg ctgaggatgc tggaggtcct gggggcttc 600

ctgcctttcc tcctgctgct cgtctgccac gtgctcaccc aggccacagc ctgtcgcacc 660

tgccaccgcc aacagcagcc cgcagcctgc cggggcttcg cccgtgtggc caggaccatt 720

ctgtcagcct atgtggtcct gaggctgccc taccagctgg cccagctgct ctacctggcc 780

ttcctgtggc acgtctactc tggctacctg ctctgggagg ccctggtcta ctccgactac 840

ctgatcctac tcaacagctg cctcagcccc ttcctctgcc tcatggccag tgccgacctc 900

cggaccctgc tgcgctccgt gctctcgtcc ttcgcggcag ctctctgcga ggagcggccg 960

ggcagcttca cgcccactga gccacagacc cagctagatt ctgagggtcc aactctgcca 1020

gagccgatgg cagaggccca gtcacagatg gatcctgtgg cccagcctca ggtgaacccc 1080

acactccagc cacgatcgga tcccacagct cagccacagc tgaaccctac ggcccagcca 1140

cagtcggatc ccacagccca gccacagctg aacctcatgg cccagccaca gtcagattct 1200

gtggcccagc cacaggcaga cactaacgtc cagacccctg cacctgctgc c 1251

<210> 68
<211> 417
<212> PRT
<213> H.Sapiens

<400> 68

```
Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro Arg Thr
1               5                   10                  15

Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp Thr Val
            20                  25                  30

Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro Ala Asn Gly Leu Met
        35                  40                  45

Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr Arg Leu
    50                  55                  60

Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe Leu Ala
65                  70                  75                  80

Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His Trp Pro
                85                  90                  95

Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly Val Ser
            100                 105                 110

Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp Arg Cys
        115                 120                 125

Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro Val Arg
    130                 135                 140

Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr Leu Phe
145                 150                 155                 160

Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp Tyr Asp
                165                 170                 175

Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser Leu Arg
            180                 185                 190

Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe Leu Leu Leu Leu Val
        195                 200                 205

Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg Thr Cys His Arg Gln
    210                 215                 220

Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg Thr Ile
225                 230                 235                 240
```

Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr Gln Leu Ala Gln Leu
245 250 255

Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu Leu Trp
260 265 270

Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser Cys Leu
275 280 285

Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr Leu Leu
290 295 300

Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Glu Glu Arg Pro
305 310 315 320

Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser Glu Gly
325 330 335

Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met Asp Pro
340 345 350

Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln Pro Arg Ser Asp Pro
355 360 365

Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser Asp Pro
370 375 380

Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser Asp Ser
385 390 395 400

Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala Pro Ala
405 410 415

Ala

<210> 69
<211> 659
<212> DNA
<213> H.Sapiens

<400> 69
tacaggcctg agcatgctgg gctccatcag caccaagcac tgcctgtcca tcctgtggcc 60

catctagtac cgctgccacc accccacaca cctgtcagca gtcgtgtgtc ctgctctggg 120

ccctgtccct gctgcagagc atcctggaat ggatgttctg tggcttcctg tctagtggtg 180

ctgattctgt ttggtgtgaa acatcagatt tcatcacagt cacatggctg atttttttat 240

gtgtggttct ctgcgggtcc agcccggttc tgctggtcag gatcctttgt ggatcccgga 300

agatgccctt gaccaggctg tacatgacca tcctgctcag agtgctggtc ttcctcctct 360

gtgacctgcc ctttggcatt cagtgattcc tatttttctg gatccacgtg gatttgtcac 420

gttcgtctag tttccatttt cctgtccact cttaacagca gtgccaaccc cattatttac 480

ttcttcatgg gctcctttag gcagcttcaa aacaggaaga ctctctagct ggttctccag 540

agggctctgc aggacacgcc tgaggtggaa gaaggcagat ggcggctttc tgaggaaacc 600

ctggagctgt catgaagcag attggggcca tgaggaagag cctctgccct gtcagtcag 659

<210> 70
<211> 213
<212> PRT
<213> H.Sapiens

<400> 70

Tyr Arg Pro Glu His Ala Gly Leu His Gln His Gln Ala Leu Pro Val
1 5 10 15

His Pro Val Ala His Leu Val Pro Leu Pro Pro Pro His Thr Pro Val
20 25 30

Ser Ser Arg Val Ser Cys Ser Gly Pro Cys Pro Cys Cys Arg Ala Ser
35 40 45

Trp Asn Gly Cys Ser Val Ala Ser Cys Leu Val Val Leu Ile Leu Phe
50 55 60

Gly Val Lys His Gln Ile Ser Ser Gln Ser His Gly Phe Phe Tyr Val
65 70 75 80

Trp Phe Ser Ala Gly Pro Ala Arg Phe Cys Trp Ser Gly Ser Phe Val
85 90 95

Asp Pro Gly Arg Cys Pro Pro Gly Cys Thr Pro Ser Cys Ser Glu Cys
100 105 110

Trp Ser Ser Ser Ser Val Thr Cys Pro Leu Ala Phe Ser Asp Ser Tyr
115 120 125

Phe Ser Gly Ser Thr Trp Ile Cys His Val Arg Leu Val Ser Ile Phe
130 135 140

Leu Ser Thr Leu Asn Ser Ser Ala Asn Pro Ile Ile Tyr Phe Phe Met
145 150 155 160

Gly Ser Phe Arg Gln Leu Gln Asn Arg Lys Thr Leu Leu Val Leu Gln
165 170 175

Arg Ala Leu Gln Asp Thr Pro Glu Val Glu Glu Gly Arg Trp Arg Leu
180 185 190

Ser Glu Glu Thr Leu Glu Leu Ser Ser Arg Leu Gly Pro Gly Arg Ala
195 200 205

Ser Ala Leu Ser Val
210

<210> 71
<211> 559
<212> DNA
<213> H.Sapiens

<400> 71
atgccgaagg caggccgcag aagagaagag gaggacggtg aggaggatga gcccagggaa 60

gccccggggt gggggccgct gggggcctcg ctccacccgc agcagcagca taaggctggc 120

cccacacatg gtgcaacaca gcagagccag cagcaccgct gccaccagcc acagcgtccg 180

gcacaagtgg cggctgggct ccccgaagaa ctgggtgcag gcgccgctga gcagcaggtg 240

cagcagcagg cagagggccc aggtgagggc gcacacacag gtggtcaggt ggcgtgggcg 300

gcggcacgag taccaggctg ggaagagggc ggccaggcac tgctccacgc tgacggccgc 360

caggagactc aggcccacga tgtagcagaa gaagcgcagc gttgccaggc tggtctgcac 420

gaagcccggg aagtccagcc ggccttgcag caagtcgggg acgatggcca ccatgtggca 480

gccaaggaag atgagatccg cgcaggccac gtccaggagg tagatggcga aagggtttct 540

gtagacattg gagctgagc 559

<210> 72
<211> 211
<212> PRT
<213> H Sapiens

<400> 72

Leu Ser Ser Asn Val Tyr Arg Asn Pro Phe Ala Ile Tyr Leu Leu Asp
1 5 10 15

Val Ala Cys Ala Asp Leu Ile Phe Leu Gly Cys His Met Val Ala Ile
20 25 30

Val Pro Asp Leu Leu Gln Gly Arg Leu Asp Phe Pro Gly Phe Val Gln
35 40 45

Thr Ser Leu Ala Thr Leu Arg Phe Phe Cys Tyr Ile Val Gly Leu Ser
50 55 60

Leu Leu Ala Ala Val Ser Val Glu Gln Cys Leu Ala Ala Leu Phe Pro
65 70 75 80

Ala Trp Tyr Ser Cys Arg Arg Pro Arg His Leu Thr Thr Cys Val Cys
85 90 95

Ala Leu Thr Trp Ala Leu Cys Leu Leu Leu His Leu Thr Thr Cys Val
100 105 110

Cys Ala Leu Thr Trp Ala Leu Cys Leu Leu Leu His Leu Leu Leu Ser
115 120 125

Gly Ala Cys Thr Leu Leu Leu Ser Gly Ala Cys Thr Gln Phe Phe Gly
130 135 140

Glu Pro Ser Arg His Leu Cys Arg Thr Leu Trp Leu Val Ala Ala Val
145 150 155 160

Leu Leu Ala Leu Leu Cys Cys Thr Met Cys Gly Ala Ser Leu Met Leu
165 170 175

Leu Leu Arg Val Glu Arg Gly Pro Gln Arg Pro Pro Pro Arg Gly Phe
180 185 190

Pro Gly Leu Ile Leu Leu Thr Val Leu Leu Phe Ser Ser Ala Ala Cys
195 200 205

Leu Arg His
210

<210> 73
<211> 1008
<212> DNA
<213> H.Sapiens

<400> 73
atggaatcat ctttctcatt tggagtgatc cttgctgtcc tggcctccct catcattgct 60

actaacacac tagtggctgt ggctgtgctg ctgttgatcc acaagaatga tggtgtcagt 120

ctctgcttca ccttgaatct ggctgtggct gacaccttga ttggtgtggc catctctggc 180

ctactcacag accagctctc cagcccttct cggcccacac agaagaccct gtgcagcctg 240

cggatggcat ttgtcacttc ctccgcagct gcctctgtcc tcacggtcat gctgatcacc 300

tttgacaggt accttgccat caagcagccc ttccgctact tgaagatcat gagtgggttc 360

gtggccgggg cctgcattgc cgggctgtgg ttagtgtctt acctcattgg cttcctccca 420

ctcggaatcc ccatgttcca gcagactgcc tacaaagggc agtgcagctt ctttgctgta 480

tttcaccctc acttcgtgct gaccctctcc tgcgttggct tcttcccagc catgctcctc 540

tttgtcttct tctactgcga catgctcaag attgcctcca tgcacagcca gcagattcga 600

aagatggaac atgcaggagc catggctgga ggttatcgat ccccacggac tcccagcgac 660

ttcaaagctc tccgtactgt gtctgttctc attgggagct ttgctctatc ctggaccccc 720

ttcctcatca ctggcattgt gcaggtggcc tgccaggagt gtcacctcta cctagtgctg 780

gaacggtacc tgtggctgct cggcgtgggc aactccctgc tcaacccact catctatgcc 840

tatttgcaga aggaggtgcg actgcagctc taccacatgg ccctaggagt gaagaaggtg 900

ctcacctcat tcctcctctt tctctcggcc aggaattgtg gcccagagag gcccagggaa 960

agttcctgtc acatcgtcac tatctccagc tcagagtttg atggctaa 1008

<210> 74
<211> 335
<212> PRT
<213> H.Sapiens

<400> 74

Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser
1   5   10   15

Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu
            20   25   30

Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala
        35   40   45

Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp
    50   55   60

Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu
65   70   75   80

Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val
                85   90   95

Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg
            100   105   110

Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly
        115   120   125

Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro
    130   135   140

Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val
145   150   155   160

Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro
                165   170   175

Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala
            180   185   190

Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met
        195   200   205

Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu
    210   215   220

Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro
225   230   235   240

Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu
                245   250   255

Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser
            260   265   270

Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu
        275   280   285

Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe
    290   295   300

Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu
305 310 315 320

Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly
325 330 335

<210> 75
<211> 2137
<212> DNA
<213> H.Sapiens

<400> 75
aactggaagg gcagccgtct gccgcccacg aacaccttct caagcacttt gagtgaccac 60

ggcttgcaag ctggtggctg gccccccgag tcccgggctc tgaggcacgg ccgtcgactt 120

aagcgttgca tcctgttacc tggagaccct ctgagctctc acctgctact tctgccgctg 180

cttctgcaca gagcccgggc gaggacccct ccaggatgca ggtcccgaac agcaccggcc 240

cggacaacgc gacgctgcag atgctgcgga acccggcgat cgcggtggcc ctgcccgtgg 300

tgtactcgct ggtggcggcg gtcagcatcc cgggcaacct cttctctctg tgggtgctgt 360

gccggcgcat ggggcccaga tccccgtcgg tcatcttcat gatcaacctg agcgtcacgg 420

acctgatgct ggccagcgtg ttgcctttcc aaatctacta ccattgcaac cgccaccact 480

gggtattcgg ggtgctgctt tgcaacgtgg tgaccgtggc cttttacgca aacatgtatt 540

ccagcatcct caccatgacc tgtatcagcg tggagcgctt cctggggtcc ctgtacccgc 600

tcagctccaa gcgctggcgc cgccgtcgtt acgcggtggc cgcgtgtgca gggacctggc 660

tgctgctcct gaccgccctg tccccgctgg cgcgcaccga tctcacctac ccggtgcacg 720

ccctgggcat catcacctgc ttcgacgtcc tcaagtggac gatgctcccc agcgtggcca 780

tgtgggccgt gttcctcttc accatcttca tcctgctgtt cctcatcccg ttcgtgatca 840

ccgtggcttg ttacacggcc accatcctca agctgttgcg cacggaggag gcgcacggcc 900

gggagcagcg gaggcgcgcg gtgggcctgg ccgcggtggt cttgctggcc tttgtcacct 960

gcttcgcccc caacaacttc gtgctcctgg cgcacatcgt gagccgcctg ttctacggca 1020

agagctacta ccacgtgtac aagctcacgc tgtgtctcag ctgcctcaac aactgtctgg 1080

acccgtttgt ttattacttt gcgtcccggg aattccagct gcgcctgcgg gaatatttgg 1140

gctgccgccg ggtgcccaga gacaccctgg acacgcgccg cgagagcctc ttctccgcca 1200

ggaccacgtc cgtgcgctcc gaggccggtg cgcaccctga agggatggag ggagccacca 1260

ggcccggcct ccagaggcag gagagtgtgt tctgagtccc cggggcgcag cttggagagc 1320

cgggggcgca gcttggagga tccaggggcg catggagagg ccacggtgcc agaggttcag 1380

ggagaacagc tgcgttgctc ccaggcactg cagaggcccg gtggggaagg gtctccaggc 1440

tttattcctc ccaggcactg cagaggcacc ggtgaggaag ggtctccagg cttcactcag 1500

ggtagagaaa caagcaaagc ccagcagcgc acagggtgct tgttatcctg cagagggtgc 1560

ctctgcctct ctgtgtcagg ggacagcttg tgtcaccacg ccccgctaat ttttgtattt 1620

tttttagtag agctgggctg tcacccccga gctccttaga cactcctcac acctgtccat 1680

acccgaggat ggatattcaa ccagccccac cgcctacccg actcggtttc tggatatcct 1740

ctgtgggcga actgcgagcc ccattcccag ctcttctccc tgctgacatc gtcccttagc 1800

acacctgtcc atacccgagg atggatattc aaccagcccc accgcctacc cgactcggtt 1860

tctggatatc ctctgtgggc gaactgcgag ccccattccc agctcttctc cctgctgaca 1920

tcgtccctta gttgtggttc tggccttctc cattctcctc caggggttct ggtctccgta 1980

gcccggtgca cgccgaaatt tctgtttatt tcactcaggg gcactgtggt tgctgtggtt 2040

ggaattcttc tttcagagga gcgcctgggg ctcctgcaag tcagctactc tccgtgccca 2100

cttcccctca cacacacacc cccctcgtgc cgaattc 2137

<210> 76
<211> 359
<212> PRT
<213> H.Sapiens

<400> 76

Met Gln Val Pro Asn Ser Thr Gly Pro Asp Asn Ala Thr Leu Gln Met
1 5 10 15

Leu Arg Asn Pro Ala Ile Ala Val Ala Leu Pro Val Val Tyr Ser Leu
20 25 30

Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Ser Leu Trp Val Leu
35 40 45

Cys Arg Arg Met Gly Pro Arg Ser Pro Ser Val Ile Phe Met Ile Asn
50 55 60

Leu Ser Val Thr Asp Leu Met Leu Ala Ser Val Leu Pro Phe Gln Ile
65 70 75 80

Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Leu Leu Cys
85 90 95

Asn Val Val Thr Val Ala Phe Tyr Ala Asn Met Tyr Ser Ser Ile Leu
100 105 110

Thr Met Thr Cys Ile Ser Val Glu Arg Phe Leu Gly Val Leu Tyr Pro
115 120 125

Leu Ser Ser Lys Arg Trp Arg Arg Arg Arg Tyr Ala Val Ala Ala Cys
130 135 140

Ala Gly Thr Trp Leu Leu Leu Leu Thr Ala Leu Ser Pro Leu Ala Arg
145 150 155 160

Thr Asp Leu Thr Tyr Pro Val His Ala Leu Gly Ile Ile Thr Cys Phe
165 170 175

Asp Val Leu Lys Trp Thr Met Leu Pro Ser Val Ala Met Trp Ala Val
180 185 190

Phe Leu Phe Thr Ile Phe Ile Leu Leu Phe Leu Ile Pro Phe Val Ile
195 200 205

Thr Val Ala Cys Tyr Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu
210 215 220

Glu Ala His Gly Arg Glu Gln Arg Arg Arg Ala Val Gly Leu Ala Ala
225 230 235 240

Val Val Leu Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val
245 250 255

Leu Leu Ala His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr
260 265 270

His Val Tyr Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu
275 280 285

Asp Pro Phe Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu
290 295 300

Arg Glu Tyr Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr
305 310 315 320

Arg Arg Glu Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu
325 330 335

Ala Gly Ala His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu
340 345 350

Gln Arg Gln Glu Ser Val Phe
355

<210> 77
<211> 1197
<212> DNA
<213> H.Sapiens

<400> 77
atggagtcgg ggctgctgcg gccggcgccg gtgagcgagg tcatcgtcct gcattacaac 60

tacaccggca agctccgcgg tgcgcgctac cagccgggtg ccggcctgcg cgccgacgcc 120

gtggtgtgcc tggcggtgtg cgccttcatc gtgctagaga atctagccgt gttgttggtg 180

ctcggacgcc acccgcgctt ccacgctccc atgttcctgc tcctgggcag cctcacgttg 240

tcggatctgc tggcaggcgc cgcctacgcc gccaacatcc tactgtcggg gccgctcacg 300

ctgaaactgt cccccgcgct ctggttcgca cgggagggag gcgtcttcgt ggcactcact 360

gcgtccgtgc tgagcctcct ggccatcgcg ctggagcgca gcctcaccat ggcgcgcagg 420

gggcccgcgc ccgtctccag tcgggggcgc acgctggcga tggcagccgc ggcctggggc 480

gtgtcgctgc tcctcgggct cctgccagcg ctgggctgga attgcctggg tcgcctggac 540

gcttgctcca ctgtcttgcc gctctacgcc aaggcctacg tgctcttctg cgtgctcgcc 600

ttcgtgggca tcctggccgc tatctgtgca ctctacgcgc gcatctactg ccaggtacgc 660

gccaacgcgc ggcgcctgcc ggcacggccc gggactgcgg ggaccacctc gacccgggcg 720

cgtcgcaagc cgcgctcgct ggccttgctg cgcacgctca gcgtggtgct cctggccttt 780

gtggcatgtt gggggcccct cttcctgctg ctgttgctcg acgtggcgtg cccggcgcgc 840

acctgtcctg tactcctgca ggccgatccc ttcctgggac tggccatggc caactcactt 900

ctgaacccca tcatctacac gctcaccaac cgcgacctgc gccacgcgct cctgcgcctg 960

gtctgctgcg gacgccactc ctgcggcaga gacccgagtg gctcccagca gtcggcgagc 1020

gcggctgagg cttccggggg cctgcgccgc tgcctgcccc cgggccttga tgggagcttc 1080

agcggctcgg agcgctcatc gccccagcgc gacgggctgg acaccagcgg ctccacaggc 1140

agccccggtg cacccacagc cgcccggact ctggtatcag aaccggctgc agactga 1197

<210> 78
<211> 398
<212> PRT
<213> H.Sapiens

<400> 78

Met Glu Ser Gly Leu Leu Arg Pro Ala Pro Val Ser Glu Val Ile Val
1 5 10 15

Leu His Tyr Asn Tyr Thr Gly Lys Leu Arg Gly Ala Arg Tyr Gln Pro
20 25 30

Gly Ala Gly Leu Arg Ala Asp Ala Val Val Cys Leu Ala Val Cys Ala
35 40 45

Phe Ile Val Leu Glu Asn Leu Ala Val Leu Leu Val Leu Gly Arg His
50 55 60

Pro Arg Phe His Ala Pro Met Phe Leu Leu Leu Gly Ser Leu Thr Leu
65 70 75 80

Ser Asp Leu Leu Ala Gly Ala Ala Tyr Ala Ala Asn Ile Leu Leu Ser
85 90 95

Gly Pro Leu Thr Leu Lys Leu Ser Pro Ala Leu Trp Phe Ala Arg Glu
100 105 110

Gly Gly Val Phe Val Ala Leu Thr Ala Ser Val Leu Ser Leu Leu Ala
115 120 125

Ile Ala Leu Glu Arg Ser Leu Thr Met Ala Arg Arg Gly Pro Ala Pro
130 135 140

Val Ser Ser Arg Gly Arg Thr Leu Ala Met Ala Ala Ala Ala Trp Gly
145 150 155 160

Val Ser Leu Leu Leu Gly Leu Leu Pro Ala Leu Gly Trp Asn Cys Leu
165 170 175

Gly Arg Leu Asp Ala Cys Ser Thr Val Leu Pro Leu Tyr Ala Lys Ala
180 185 190

Tyr Val Leu Phe Cys Val Leu Ala Phe Val Gly Ile Leu Ala Ala Ile
195 200 205

Cys Ala Leu Tyr Ala Arg Ile Tyr Cys Gln Val Arg Ala Asn Ala Arg
210 215 220

Arg Leu Pro Ala Arg Pro Gly Thr Ala Gly Thr Thr Ser Thr Arg Ala
225 230 235 240

Arg Arg Lys Pro Arg Ser Leu Ala Leu Leu Arg Thr Leu Ser Val Val
245 250 255

Leu Leu Ala Phe Val Ala Cys Trp Gly Pro Leu Phe Leu Leu Leu Leu
260 265 270

Leu Asp Val Ala Cys Pro Ala Arg Thr Cys Pro Val Leu Leu Gln Ala
275 280 285

Asp Pro Phe Leu Gly Leu Ala Met Ala Asn Ser Leu Leu Asn Pro Ile
290 295 300

Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg His Ala Leu Leu Arg Leu
305 310 315 320

Val Cys Cys Gly Arg His Ser Cys Gly Arg Asp Pro Ser Gly Ser Gln
325 330 335

Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly Gly Leu Arg Arg Cys Leu
340 345 350

Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly Ser Glu Arg Ser Ser Pro
355 360 365

Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser Thr Gly Ser Pro Gly Ala
370 375 380

Pro Thr Ala Ala Arg Thr Leu Val Ser Glu Pro Ala Ala Asp
385 390 395

<210> 79
<211> 1041
<212> DNA
<213> H.Sapiens

<400> 79
atgtacaacg ggtcgtgctg ccgcatcgag ggggacacca tctcccaggt gatgccgccg 60

ctgctcattg tggcctttgt gctgggcgca ctaggcaatg ggtcgccct gtgtggtttc 120

tgcttccaca tgaagacctg gaagcccagc actgtttacc ttttcaattt ggccgtggct 180

gatttcctcc ttatgatctg cctgcctttt cggacagact attacctcag acgtagacac 240

tgggcttttg gggacattcc ctgccgagtg ggctcttca cgttggccat gaacagggcc 300

gggagcatcg tgttccttac ggtggtggct gcggacaggt atttcaaagt ggtccacccc 360

caccacgcgg tgaacactat ctccacccgg gtggcggctg gcatcgtctg caccctgtgg 420

gccctggtca tcctgggaac agtgtatctt ttgctggaga accatctctg cgtgcaagag 480

acggccgtct cctgtgagag cttcatcatg gagtcggcca atggctggca tgacatcatg 540

ttccagctgg agttctttat gcccctcggc atcatcttat tttgctcctt caagattgtt 600

tggagcctga ggcggaggca gcagctggcc agacaggctc ggatgaagaa ggcgacccgg 660

ttcatcatgg tggtggcaat tgtgttcatc acatgctacc tgcccagcgt gtctgctaga 720

ctctatttcc tctggacggt gccctcgagt gcctgcgatc cctctgtcca tggggccctg 780

cacataaccc tcagcttcac ctacatgaac agcatgctgg atcccctggt gtattatttt 840

tcaagcccct cctttcccaa attctacaac aagctcaaaa tctgcagtct gaaacccaag 900

cagccaggac actcaaaaac acaaaggccg gaagagatgc caatttcgaa cctcggtcgc 960

aggagttgca tcagtgtggc aaatagtttc caaagccagt ctgatggcaa atgggatccc 1020

cacattgttg agtggcactg a 1041

<210> 80
<211> 346
<212> PRT
<213> H.Sapiens

<400> 80

Met Tyr Asn Gly Ser Cys Cys Arg Ile Glu Gly Asp Thr Ile Ser Gln
1 5 10 15

Val Met Pro Pro Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly
20 25 30

Asn Gly Val Ala Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys
35 40 45

Pro Ser Thr Val Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu
50 55 60

Met Ile Cys Leu Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His
65 70 75 80

Trp Ala Phe Gly Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala
85 90 95

Met Asn Arg Ala Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp
100 105 110

Arg Tyr Phe Lys Val Val His Pro His His Ala Val Asn Thr Ile Ser
115 120 125

Thr Arg Val Ala Ala Gly Ile Val Cys Thr Leu Trp Ala Leu Val Ile
130 135 140

Leu Gly Thr Val Tyr Leu Leu Leu Glu Asn His Leu Cys Val Gln Glu
145 150 155 160

Thr Ala Val Ser Cys Glu Ser Phe Ile Met Glu Ser Ala Asn Gly Trp
165 170 175

His Asp Ile Met Phe Gln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile
180 185 190

Leu Phe Cys Ser Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln
195 200 205

Leu Ala Arg Gln Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val
210 215 220

Val Ala Ile Val Phe Ile Thr Cys Tyr Leu Pro Ser Val Ser Ala Arg
225 230 235 240

Leu Tyr Phe Leu Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val
245 250 255

His Gly Ala Leu His Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met
260 265 270

Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe
275 280 285

Tyr Asn Lys Leu Lys Ile Cys Ser Leu Lys Pro Lys Gln Pro Gly His
290 295 300

Ser Lys Thr Gln Arg Pro Glu Glu Met Pro Ile Ser Asn Leu Gly Arg
305 310 315 320

Arg Ser Cys Ile Ser Val Ala Asn Ser Phe Gln Ser Gln Ser Asp Gly
325 330 335

Gln Trp Asp Pro His Ile Val Glu Trp His
340 345

<210> 81
<211> 2525
<212> DNA
<213> H.Sapiens

<400> 81
caagaatgac aggtgacttc ccaagtatgc ctggccacaa tacctccagg aattcctctt 60

gcgatcctat agtgacaccc cacttaatca gcctctactt catagtgctt attggcgggc 120

tggtgggtgt catttccatt cttttcctcc tggtgaaaat gaacacccgg tcagtgacca 180

ccatggcggt cattaacttg gtggtggtcc acagcgtttt tctgctgaca gtgccatttc 240

gcttgaccta cctcatcaag aagacttgga tgtttgggct gcccttctgc aaatttgtga 300

gtgccatgct gcacatccac atgtacctca cgttcctatt ctatgtggtg atcctggtca 360

ccagatacct catcttcttc aagtgcaaag acaaagtgga attctacaga aaactgcatg 420

ctgtggctgc cagtgctggc atgtggacgc tggtgattgt cattgtggta cccctggttg 480

tctcccggta tggaatccat gaggaataca atgaggagca ctgttttaaa tttcacaaag 540

agcttgctta cacatatgtg aaaatcatca actatatgat agtcattttt gtcatagccg 600

ttgctgtgat tctgttggtc ttccaggtct tcatcattat gttgatggtg cagaagctac 660

gccactcttt actatcccac caggagttct gggctcagct gaaaaaccta ttttttatag 720

gggtcatcct tgtttgtttc cttccctacc agttctttag gatctattac ttgaatgttg 780

tgacgcattc caatgcctgt aacagcaagg ttgcatttta taacgaaatc ttcttgagtg 840

taacagcaat tagctgctat gatttgcttc tctttgtctt tggggaagc cattggttta 900

agcaaaagat aattggctta tggaattgtg ttttgtgccg ttagccacaa actacagtat 960

tcatatttgc ttcctttata ttggaataa aaatgggtat aggggaggta agaatggtat 1020

ttcattactt gatcaaaacc atgccttgat gtacccaaaa caaaaggact ataaaatgca 1080

agagccctca ttgtagtcct tatgggatcc ctcccatctc tgagtgatgg ccgtacaaag 1140

accagtgttg ttgaatccac ctggagttgc aatattacat tattttccag tacagaatgt 1200

ctgtgtggcc catgaaagca acataggttt taagagtttt agagtttcat tagctcattc 1260

taagttcctc tgtttgaagc atggtctctt aggttttgga ctgaactcag acctttagtt 1320

cttttcatcc cacttcacct taggtaagta aattctggcc accacccagc tccaaagaca 1380

caaactctcc ttcgctaacc aggttagatg tcccattcat ctcatgccct gataaaaact 1440

gataaggga gagaatagtt aaaaattttt ctagggtatc ataactctgg taggaagtca 1500

tctgtctaga aatcaagaga aaaagaacgt gtggcctcct gttataacaa gggtttctag 1560

atttgtcctg tgaaaggtcg tttaaggact tggggatcaa cttcctcaat tatcaccaat 1620

tgcactgttg ctccaaaaat catttaaaag cttactggac atatctacat aatggtgaaa 1680

ctgtaattta gagactatcc ctgactaatg tgctggtagg cattaaaatg agttcccaag 1740

ggaagtgatt aaaatttttt tctcttctgt tttttgagag aatttctaga tgtcctgggc 1800

cacagttaat taagattttt agggggaca gaaagttata ctgaaatctt tagagctccc 1860

ttccgccgtt aaaattatat atatatatat ttaaattata ccttaagttc tggggtacat 1920

gtgcagaatg tgcaggtttg ttacataggt atacacgtgc catggtggtt tgcggcacct 1980

gtcaacccat ctacattagg tatttctcct aatgctctcc ctcccctagc cccccacccc 2040

tggacaggcc ccattgtgtg atgttcccct ccctgtgtcc atgtgttttc attgttcaac 2100

tcccacttct aagtgagaac atgcggtgtt tggttttctg ttcctgtgtt agtttgctga 2160

gaatgatggt ttccaggtta aaattatata tttttaaata aatgaaaact gtgtttttaa 2220

aagaggactt ttgagaagta tatagaaaaa ccattaattt agactctgtg agattaggtt 2280

gcatgaagaa ggttttctga atatttgaag agtggataaa taaatgtccc ccaaagcaat 2340

aaaatcataa tcctttaaaa tataggaaaa ataactaatg ggaactaggc ttaatactcg 2400

ggatgaaata atctgtacaa caaactccca tgacacatgt ttacctatgt aacaaacctg 2460

cacatgtacc cctgaactta aaataaaatt taaagtataa taataaaata atatggattt 2520

tcttt 2525

<210> 82
<211> 312
<212> PRT
<213> H.Sapiens

<400> 82

Met Thr Gly Asp Phe Pro Ser Met Pro Gly His Asn Thr Ser Arg Asn
1 5 10 15

Ser Ser Cys Asp Pro Ile Val Thr Pro His Leu Ile Ser Leu Tyr Phe
20 25 30

Ile Val Leu Ile Gly Gly Leu Val Gly Val Ile Ser Ile Leu Phe Leu
35 40 45

Leu Val Lys Met Asn Thr Arg Ser Val Thr Thr Met Ala Val Ile Asn
50 55 60

Leu Val Val Val His Ser Val Phe Leu Leu Thr Val Pro Phe Arg Leu
65 70 75 80

Thr Tyr Leu Ile Lys Lys Thr Trp Met Phe Gly Leu Pro Phe Cys Lys
85 90 95

Phe Val Ser Ala Met Leu His Ile His Met Tyr Leu Thr Phe Leu Phe
100 105 110

Tyr Val Val Ile Leu Val Thr Arg Tyr Leu Ile Phe Phe Lys Cys Lys
115 120 125

Asp Lys Val Glu Phe Tyr Arg Lys Leu His Ala Val Ala Ala Ser Ala

130 135 140

Gly Met Trp Thr Leu Val Ile Val Ile Val Val Pro Leu Val Val Ser
145 150 155 160

Arg Tyr Gly Ile His Glu Glu Tyr Asn Glu Glu His Cys Phe Lys Phe
165 170 175

His Lys Glu Leu Ala Tyr Thr Tyr Val Lys Ile Ile Asn Tyr Met Ile
180 185 190

Val Ile Phe Val Ile Ala Val Ala Val Ile Leu Leu Val Phe Gln Val
195 200 205

Phe Ile Ile Met Leu Met Val Gln Lys Leu Arg His Ser Leu Leu Ser
210 215 220

His Gln Glu Phe Trp Ala Gln Leu Lys Asn Leu Phe Phe Ile Gly Val
225 230 235 240

Ile Leu Val Cys Phe Leu Pro Tyr Gln Phe Phe Arg Ile Tyr Tyr Leu
245 250 255

Asn Val Val Thr His Ser Asn Ala Cys Asn Ser Lys Val Ala Phe Tyr
260 265 270

Asn Glu Ile Phe Leu Ser Val Thr Ala Ile Ser Cys Tyr Asp Leu Leu
275 280 285

Leu Phe Val Phe Gly Gly Ser His Trp Phe Lys Gln Lys Ile Ile Gly
290 295 300

Leu Trp Asn Cys Val Leu Cys Arg
305 310

<210> 83
<211> 1125
<212> DNA
<213> H.Sapiens

<400> 83
gcaggagcac tgaaaatcag gaacaatcct gtatttttg tgataatcaa caaggacaaa 60

acttctccat atgtaaataa cagcgttatg agcagcaatt catccctgct ggtggctgtg 120

cagctgtgct acgcgaacgt gaatgggtcc tgtgtgaaaa tcccctctc gccgggatcc 180

cgggtgattc tgtacatagt gtttggcttt ggggctgtgc tggctgtgtt tggaaacctc 240

ctggtgatga tttcaatcct ccatttcaag cagctgcact ctccgaccaa ttttctcgtt 300

gcctctctgg cctgcgctga tttcttggtg ggtgtgactg tgatgccctt cagcatggtc 360

aggacggtgg agagctgctg gtattttggg aggagttttt gtactttcca cacctgctgt 420

gatgtggcat tttgttactc ttctctcttt cacttgtgct tcatctccat cgacaggtac 480

attgcggtta ctgaccccct ggtctatcct accaagttca ccgtatctgt gtcaggaatt 540

tgcatcagcg tgtcctggat cctgcccctc atgtacagcg gtgctgtgtt ctacacaggt 600

gtctatgacg atgggctgga ggaattatct gatgccctaa actgtatagg aggttgtcag 660

accgttgtaa atcaaaactg ggtgttgaca gattttctat tcttctttat acctaccttt 720

attatgataa ttctgtatgg taacatattt cttgtggcta gacgacaggc gaaaaagata 780

gaaaatactg gtagcaagac agaatcatcc tcagagagtt acaaagccag agtggccagg 840

agagagagaa aagcagctaa aaccctgggg gtcacagtgg tagcatttat gatttcatgg 900

ttaccatata gcattgattc attaattgat gcctttatgg gctttataac ccctgcctgt 960

atttatgaga tttgctgttg gtgtgcttat tataactcag ccatgaatcc tttgatttat 1020

gctttatttt acccatggtt taggaaagca ataaaagtta ttgtaactgg tcaggtttta 1080

aagaacagtt cagcaaccat gaatttgttt tctgaacata tataa 1125

<210> 84
<211> 345
<212> PRT
<213> H.Sapiens

<400> 84

Met Ser Ser Asn Ser Ser Leu Leu Val Ala Val Gln Leu Cys Tyr Ala
1 5 10 15

Asn Val Asn Gly Ser Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg
20 25 30

Val Ile Leu Tyr Ile Val Phe Gly Phe Gly Ala Val Leu Ala Val Phe
35 40 45

Gly Asn Leu Leu Val Met Ile Ser Ile Leu His Phe Lys Gln Leu His
50 55 60

Ser Pro Thr Asn Phe Leu Val Ala Ser Leu Ala Cys Ala Asp Phe Leu
65 70 75 80

Val Gly Val Thr Val Met Pro Phe Ser Met Val Arg Thr Val Glu Ser
85 90 95

Cys Trp Tyr Phe Gly Arg Ser Phe Cys Thr Phe His Thr Cys Cys Asp
100 105 110

Val Ala Phe Cys Tyr Ser Ser Leu Phe His Leu Cys Phe Ile Ser Ile
115 120 125

Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Val Tyr Pro Thr Lys Phe
130 135 140

Thr Val Ser Val Ser Gly Ile Cys Ile Ser Val Ser Trp Ile Leu Pro
145 150 155 160

Leu Met Tyr Ser Gly Ala Val Phe Tyr Thr Gly Val Tyr Asp Asp Gly

```
                165                 170                 175

Leu Glu Glu Leu Ser Asp Ala Leu Asn Cys Ile Gly Gly Cys Gln Thr
            180                 185                 190

Val Val Asn Gln Asn Trp Val Leu Thr Asp Phe Leu Ser Phe Phe Ile
        195                 200                 205

Pro Thr Phe Ile Met Ile Ile Leu Tyr Gly Asn Ile Phe Leu Val Ala
    210                 215                 220

Arg Arg Gln Ala Lys Lys Ile Glu Asn Thr Gly Ser Lys Thr Glu Ser
225                 230                 235                 240

Ser Ser Glu Ser Tyr Lys Ala Arg Val Ala Arg Arg Glu Arg Lys Ala
                245                 250                 255

Ala Lys Thr Leu Gly Val Thr Val Val Ala Phe Met Ile Ser Trp Leu
            260                 265                 270

Pro Tyr Ser Ile Asp Ser Leu Ile Asp Ala Phe Met Gly Phe Ile Thr
        275                 280                 285

Pro Ala Cys Ile Tyr Glu Ile Cys Cys Trp Cys Ala Tyr Tyr Asn Ser
    290                 295                 300

Ala Met Asn Pro Leu Ile Tyr Ala Leu Phe Tyr Pro Trp Phe Arg Lys
305                 310                 315                 320

Ala Ile Lys Val Ile Val Thr Gly Gln Val Leu Lys Asn Ser Ser Ala
                325                 330                 335

Thr Met Asn Leu Phe Ser Glu His Ile
            340                 345

<210>  85
<211>  1020
<212>  DNA
<213>  H.Sapiens

<400>  85
accatgaatg agccactaga ctatttagca aatgcttctg atttccccga ttatgcagct      60

gcttttggaa attgcactga tgaaaacatc ccactcaaga tgcactacct ccctgttatt     120

tatggcatta tcttcctcgt gggatttcca ggcaatgcag tagtgatatc cacttacatt     180

ttcaaaatga gaccttggaa gagcagcacc atcattatgc tgaacctggc ctgcacagat     240

ctgctgtatc tgaccagcct ccccttcctg attcactact atgccagtgg cgaaaactgg     300

atctttggag atttcatgtg taagtttatc cgcttcagct tccatttcaa cctgtatagc     360

agcatcctct tcctcacctg tttcagcatc ttccgctact gtgtgatcat tcacccaatg     420

agctgctttt ccattcacaa aactcgatgt gcagttgtag cctgtgctgt ggtgtggatc     480

atttcactgg tagctgtcat tccgatgacc ttcttgatca catcaaccaa caggaccaac     540
```

agatcagcct gtctcgacct caccagttcg gatgaactca atactattaa gtggtacaac 600

ctgattttga ctgcaagtac tttctgcctc cccttggtga tagtgacact ttgctatacc 660

acgattatcc acactttgac ccatggactg caaactgaca gctgccttaa gcagaaagca 720

cgaaggctaa ccattctgct actccttgca ttttacgtat gtttttacc cttccatatc 780

ttgagggtca ttcaggatcg aatctcaccc tgctttcaat cagttgttcc attgagaatc 840

agatccatga agcttacatc gtttctagac cattatgctg ctctgaacac ctttggtaac 900

ctgttactat atgtggtggt cagcgacaac tttcagcagg ctgtctgctc aacagtgaga 960

tgcaaagtaa gcgggaacct tgagcaagca aagaaaatta gttactcaaa caacccttga 1020

<210> 86
<211> 336
<212> PRT
<213> H.Sapiens

<400> 86

Met Asn Glu Pro Leu Asp Tyr Leu Ala Asn Ala Ser Asp Phe Pro Asp
1 5 10 15

Tyr Ala Ala Ala Phe Gly Asn Cys Thr Asp Glu Asn Ile Pro Leu Lys
20 25 30

Met His Tyr Leu Pro Val Ile Tyr Gly Ile Ile Phe Leu Val Gly Phe
35 40 45

Pro Gly Asn Ala Val Val Ile Ser Thr Tyr Ile Phe Lys Met Arg Pro
50 55 60

Trp Lys Ser Ser Thr Ile Ile Met Leu Asn Leu Ala Cys Thr Asp Leu
65 70 75 80

Leu Tyr Leu Thr Ser Leu Pro Phe Leu Ile His Tyr Tyr Ala Ser Gly
85 90 95

Glu Asn Trp Ile Phe Gly Asp Phe Met Cys Lys Phe Ile Arg Phe Ser
100 105 110

Phe His Phe Asn Leu Tyr Ser Ser Ile Leu Phe Leu Thr Cys Phe Ser
115 120 125

Ile Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile
130 135 140

His Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile
145 150 155 160

Ser Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Thr Asn
165 170 175

Arg Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu
180 185 190

```
Asn Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Ser Thr Phe Cys
        195                 200                 205

Leu Pro Leu Val Ile Val Thr Leu Cys Tyr Thr Thr Ile Ile His Thr
    210                 215                 220

Leu Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ala Arg
225                 230                 235                 240

Arg Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro
                245                 250                 255

Phe His Ile Leu Arg Val Ile Gln Asp Arg Ile Ser Ala Cys Phe Gln
            260                 265                 270

Ser Val Val Pro Leu Arg Ile Arg Ser Met Lys Leu Thr Ser Phe Leu
        275                 280                 285

Asp His Tyr Ala Ala Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val
    290                 295                 300

Val Val Ser Asp Asn Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys
305                 310                 315                 320

Lys Val Ser Gly Asn Leu Glu Gln Ala Lys Lys Ile Ser Tyr Ser Asn
                325                 330                 335

<210>  87
<211>  1138
<212>  DNA
<213>  H.Sapiens

<400>  87
aaaaattgct gtactgaact attgaatgga acttggaaat aaagtccctt ccaaaataac      60

tattcttcaa cagagagtaa taggtaaatg ttttagaagt gagaggactc aaattgccaa     120

tgatttactc ttttattttt cctcctaggt ttctgggata agtatgtgca aataaaaaat     180

aaacatgaga aggaactgta acctgattat ggatttggga aaaagataaa tcaacacaca     240

aagggaaaag taaactgatt gacagccctc aggaatgatg ccctttgccc acaatataat     300

taatatttcc tgtgtgaaaa acaactggtc aaatgatgtc cgtgcttccc tgtacagttt     360

aatggtgctc ataattctga ccacactcgt tggcaatctg atagttattg tttctatatc     420

acacttcaaa caacttcata ccccaacaaa ttggctcatt cattccatgg ccactgtgga     480

ctttcttctg ggtgtctgg tcatgcctta cagtatggtg agatctgctg agcactgttg     540

gtattttgga gaagtcttct gtaaattca cacaagcacc gacattatgc tgagctcagc     600

ctccattttc catttgtctt tcatctccat tgaccgctac tatgctgtgt gtgatccact     660

gagatataaa gccaagatga atatcttggt tatttgtgtg atgatcttca ttagttggag     720

tgtccctgct gtttttgcat ttggaatgat ctttctggag ctaaacttca aaggcgctga     780
```

agagatatat tacaaacatg ttcactgcag aggaggttgc tctgtcttct ttagcaaaat 840

atctggggta ctgaccttta tgacttcttt ttatatacct ggatctatta tgttatgtgt 900

ctattacaga atatatctta tcgctaaaga acaggcaaga ttaattagtg atgccaatca 960

gaagctccaa attggattgg aaatgaaaaa tggaatttca caaagcaaag aaaggaaagc 1020

tgtgaagaca ttgggattg tgatgggagt tttcctaata tgctggtgcc ctttctttat 1080

ctgtacagtc atggaccctt ttcttcacta cattattcca cctactttga atgatgta 1138

<210> 88
<211> 296
<212> PRT
<213> H.Sapiens

<400> 88

Met Met Pro Phe Cys His Asn Ile Ile Asn Ile Ser Cys Val Lys Asn
1 5 10 15

Asn Trp Ser Asn Asp Val Arg Ala Ser Leu Tyr Ser Leu Met Val Leu
20 25 30

Ile Ile Leu Thr Thr Leu Val Gly Asn Leu Ile Val Ile Val Ser Ile
35 40 45

Ser His Phe Lys Gln Leu His Thr Pro Thr Asn Trp Leu Ile His Ser
50 55 60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser
65 70 75 80

Met Val Arg Ser Ala Glu His Cys Trp Tyr Phe Gly Glu Val Phe Cys
85 90 95

Lys Ile His Thr Ser Thr Asp Ile Met Leu Ser Ser Ala Ser Ile Phe
100 105 110

His Leu Ser Phe Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro
115 120 125

Leu Arg Tyr Lys Ala Lys Met Asn Ile Leu Val Ile Cys Val Met Ile
130 135 140

Phe Ile Ser Trp Ser Val Pro Ala Val Phe Ala Phe Gly Met Ile Phe
145 150 155 160

Leu Glu Leu Asn Phe Lys Gly Ala Glu Glu Ile Tyr Tyr Lys His Val
165 170 175

His Cys Arg Gly Gly Cys Ser Val Phe Phe Ser Lys Ile Ser Gly Val
180 185 190

Leu Thr Phe Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Leu Cys
195 200 205

```
Val Tyr Tyr Arg Ile Tyr Leu Ile Ala Lys Glu Gln Ala Arg Leu Ile
    210                 215                 220

Ser Asp Ala Asn Gln Lys Leu Gln Ile Gly Leu Glu Met Lys Asn Gly
225                 230                 235                 240

Ile Ser Gln Ser Lys Glu Arg Lys Ala Val Lys Thr Leu Gly Ile Val
                245                 250                 255

Met Gly Val Phe Leu Ile Cys Trp Cys Pro Phe Phe Ile Cys Thr Val
            260                 265                 270

Met Asp Pro Phe Leu His Tyr Ile Ile Pro Pro Thr Leu Asn Asp Ala
        275                 280                 285

Arg Gly Ser Arg Ala Asn Ser Ala
    290                 295
```

<210> 89
<211> 1023
<212> DNA
<213> H.Sapiens

<400> 89

```
ggaatgatgc cttttgcca caatataatt aatatttcct gtgtgaaaaa caactggtca      60

aatgatgtcc gtgcttccct gtacagttta atggtgctca taattctgac cacactcgtt     120

ggcaatctga tagttattgt tttatatca cacttcaaac aacttcatac cccaacaaat     180

tggctcattc attccatggc cactgtggac tttcttctgg ggtgtctggt catgccttac     240

agtatggtga gatctgctga gcactgttgg tattttggag aagtcttctg taaaattcac     300

acaagcaccg acattatgct gagctcagcc tccattttcc atttgtcttt catctccatt     360

gaccgctact atgctgtgtg tgatccactg agatataaag ccaagatgaa tatcttggtt     420

atttgtgtga tgatcttcat tagttggagt gtccctgctg tttttgcatt tggaatgatc     480

tttctggagc taaacttcaa aggcgctgaa gagatatatt acaaacatgt tcactgcaga     540

ggaggttgct ctgtcttctt tagcaaaata tctggggtac tgacctttat gacttctttt     600

tatatacctg gatctattat gttatgtgtc tattacagaa tatatcttat cgctaaagaa     660

caggcaagat taattagtga tgccaatcag aagctccaaa ttggattgga aatgaaaaat     720

ggaatttcac aaagcaaaga aaggaaagct gtgaagacat tggggattgt gatgggagtt     780

ttcctaatat gctggtgccc tttctttatc tgtacagtca tggacccttt tcttcactac     840

attattccac ctactttgaa tgatgtattg atttggtttg gctacttgaa ctctacattt     900

aatccaatgg tttatgcatt tttctatcct tggtttagaa aagcactgaa gatgatgctg     960

tttggtaaaa ttttccaaaa agattcatcc aggtgtaaat tatttttgga attgagttca    1020
```

tag 1023

<210> 90
<211> 339
<212> PRT
<213> H.Sapiens

<400> 90

Met Met Pro Phe Cys His Asn Ile Ile Asn Ile Ser Cys Val Lys Asn
1   5   10   15

Asn Trp Ser Asn Asp Val Arg Ala Ser Leu Tyr Ser Leu Met Val Leu
            20   25   30

Ile Ile Leu Thr Thr Leu Val Gly Asn Leu Ile Val Ile Val Ser Ile
        35   40   45

Ser His Phe Lys Gln Leu His Thr Pro Thr Asn Trp Leu Ile His Ser
    50   55   60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser
65   70   75   80

Met Val Arg Ser Ala Glu His Cys Trp Tyr Phe Gly Glu Val Phe Cys
                85   90   95

Lys Ile His Thr Ser Thr Asp Ile Met Leu Ser Ser Ala Ser Ile Phe
            100   105   110

His Leu Ser Phe Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro
        115   120   125

Leu Arg Tyr Lys Ala Lys Met Asn Ile Leu Val Ile Cys Val Met Ile
    130   135   140

Phe Ile Ser Trp Ser Val Pro Ala Val Phe Ala Phe Gly Met Ile Phe
145   150   155   160

Leu Glu Leu Asn Phe Lys Gly Ala Glu Glu Ile Tyr Tyr Lys His Val
                165   170   175

His Cys Arg Gly Gly Cys Ser Val Phe Phe Ser Lys Ile Ser Gly Val
            180   185   190

Leu Thr Phe Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Leu Cys
        195   200   205

Val Tyr Tyr Arg Ile Tyr Leu Ile Ala Lys Glu Gln Ala Arg Leu Ile
    210   215   220

Ser Asp Ala Asn Gln Lys Leu Gln Ile Gly Leu Glu Met Lys Asn Gly
225   230   235   240

Ile Ser Gln Ser Lys Glu Arg Lys Ala Val Lys Thr Leu Gly Ile Val
                245   250   255

Met Gly Val Phe Leu Ile Cys Trp Cys Pro Phe Phe Ile Cys Thr Val

```
            260                 265                 270

Met Asp Pro Phe Leu His Tyr Ile Ile Pro Pro Thr Leu Asn Asp Val
        275                 280                 285

Leu Ile Trp Phe Gly Tyr Leu Asn Ser Thr Phe Asn Pro Met Val Tyr
    290                 295                 300

Ala Phe Phe Tyr Pro Trp Phe Arg Lys Ala Leu Lys Met Met Leu Phe
305                 310                 315                 320

Gly Lys Ile Phe Gln Lys Asp Ser Ser Arg Cys Lys Leu Phe Leu Glu
                325                 330                 335

Leu Ser Ser


<210>  91
<211>  1696
<212>  DNA
<213>  H.Sapiens

<400>  91
ctgtaaagta gattgtatga ggactccatg aggtcatcca cttcaagtcc ttggcatagg      60

ataattactc aaaaggtgat gacaatggcg cagggaggga tggtgacttg cctggagatg     120

cacagcaccg tctctcccat actcggtcat tcacaccatc attgattcac caggcaccac     180

tccgtgtcca gcaggactct ggggacccca aatggacact accatggaag ctgacctggg     240

tgccactggc cacaggcccc gcacagagct tgatgatgag gactcctacc cccaaggtgg     300

ctgggacacg gtcttcctgg tggccctgct gctccttggg ctgccagcca atgggttgat     360

ggcgtggctg gccggctccc aggcccggca tggagctggc acgcgtctgg cgctgctcct     420

gctcagcctg gccctctctg acttcttgtt cctggcagca gcggccttcc agatcctaga     480

gatccggcat gggggacact ggccgctggg gacagctgcc tgccgcttct actacttcct     540

atggggcgtg tcctactcct ccggcctctt cctgctggcc gccctcagcc tcgaccgctg     600

cctgctggcg ctgtgcccac actggtaccc tgggcaccgc ccagtccgcc tgcccctctg     660

ggtctgcgcc ggtgtctggg tgctggccac actcttcagc gtgccctggc tggtcttccc     720

cgaggctgcc gtctggtggt acgacctggt catctgcctg gacttctggg acagcgagga     780

gctgtcgctg aggatgctgg aggtcctggg gggcttcctg cctttcctcc tgctgctcgt     840

ctgccacgtg ctcacccagg ccacagcctg tcgcacctgc caccgccaac agcagcccgc     900

agcctgccgg ggcttcgccc gtgtggccag gaccattctg tcagcctatg tggtcctgag     960

gctgccctac cagctggccc agctgctcta cctggccttc ctgtgggacg tctactctgg    1020

ctacctgctc tgggaggccc tggtctactc cgactacctg atcctactca acagctgcct    1080
```

cagccccttc ctctgcctca tggccagtgc cgacctccgg accctgctgc gctccgtgct 1140

ctcgtccttc gcggcagctc tctgcgagga gcggccgggc agcttcacgc ccactgagcc 1200

acagacccag ctagattctg agggtccaac tctgccagag ccgatggcag aggcccagtc 1260

acagatggat cctgtggccc agcctcaggt gaaccccaca ctccagccac gatcggatcc 1320

cacagctcag ccacagctga accctacggc ccagccacag tcggatccca cagcccagcc 1380

acagctgaac ctcatggccc agccacagtc agattctgtg gcccagccac aggcagacac 1440

taacgtccag acccctgcac ctgctgccag ttctgtgccc agtccctgtg atgaagcttc 1500

cccaacccca tcctcgcatc ctaccccagg ggcccttgag gacccagcca cacctcctgc 1560

ctctgaagga gaaagcccca gcagcacccc gccagaggcg gccccgggcg caggccccac 1620

gtgagggtcc aggaacacgc aggcccacca gagcagtgaa agagcccagg gcagacagag 1680

gaaccagcca gtcaga 1696

<210> 92
<211> 505
<212> PRT
<213> H.Sapiens

<400> 92

Leu Ala Trp Arg Cys Thr Ala Pro Ser Leu Pro Tyr Ser Val Ile His
1 5 10 15

Thr Ile Ile Asp Ser Pro Gly Thr Thr Pro Cys Pro Ala Gly Leu Trp
20 25 30

Gly Pro Gln Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly
35 40 45

His Arg Pro Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly
50 55 60

Gly Trp Asp Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro
65 70 75 80

Ala Asn Gly Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly
85 90 95

Ala Gly Thr Arg Leu Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp
100 105 110

Phe Leu Phe Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His
115 120 125

Gly Gly His Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe
130 135 140

Leu Trp Gly Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu
145 150 155 160

```
Ser Leu Asp Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly
                165                 170                 175

His Arg Pro Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val
            180                 185                 190

Leu Ala Thr Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala
        195                 200                 205

Val Trp Trp Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu
    210                 215                 220

Glu Leu Ser Leu Arg Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe
225                 230                 235                 240

Leu Leu Leu Leu Val Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg
                245                 250                 255

Thr Cys His Arg Gln Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg
            260                 265                 270

Val Ala Arg Thr Ile Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr
        275                 280                 285

Gln Leu Ala Gln Leu Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser
    290                 295                 300

Gly Tyr Leu Leu Trp Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu
305                 310                 315                 320

Leu Asn Ser Cys Leu Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp
                325                 330                 335

Leu Arg Thr Leu Leu Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu
            340                 345                 350

Cys Glu Glu Arg Pro Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln
        355                 360                 365

Leu Asp Ser Glu Gly Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln
    370                 375                 380

Ser Gln Met Asp Pro Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln
385                 390                 395                 400

Pro Arg Ser Asp Pro Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln
                405                 410                 415

Pro Gln Ser Asp Pro Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln
            420                 425                 430

Pro Gln Ser Asp Ser Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln
        435                 440                 445

Thr Pro Ala Pro Ala Ala Ser Ser Val Pro Ser Pro Cys Asp Glu Ala
    450                 455                 460

Ser Pro Thr Pro Ser Ser His Pro Thr Pro Gly Ala Leu Glu Asp Pro
```

```
465                 470                 475                 480

Ala Thr Pro Pro Ala Ser Glu Gly Glu Ser Pro Ser Ser Thr Pro Pro
                485                 490                 495

Glu Ala Ala Pro Gly Ala Gly Pro Thr
            500                 505

<210> 93
<211> 1413
<212> DNA
<213> H.Sapiens

<400> 93
atggacacta ccatggaagc tgacctgggt gccactggcc acaggccccg cacagagctt      60

gatgatgagg actcctaccc ccaaggtggc tgggacacgg tcttcctggt ggccctgctg     120

ctccttgggc tgccagccaa tgggttgatg gcgtggctgg ccggctccca ggcccggcat     180

ggagctggca cgcgtctggc gctgctcctg ctcagcctgg ccctctctga cttcttgttc     240

ctggcagcag cggccttcca gatcctagag atccggcatg gggacactgg gccgctgggg     300

acagctgcct gccgcttcta ctacttccta tggggcgtgt cctactcctc cggcctcttc     360

ctgctggctg ccctcagcct cgaccgctgc ctgctggcgc tgtgcccaca ctggtaccct     420

gggcaccgcc cagtccgcct gcccctctgg gtctgcgccg gtgtctgggt gctggccaca     480

ctcttcagcg tgccctggct ggtcttcccc gaggctgccg tctggtggta cgacctggtc     540

atctgcctgg acttctggga cagcgaggag ctgtcgctga ggatgctgga ggtcctgggg     600

ggcttcctgc ctttcctcct gctgctcgtc tgccacgtgc tcacccaggc cacagcctgt     660

cgcacctgcc accgccaaca gcagcccgca gcctgccggg gcttcgcccg tgtggccagg     720

accattctgt cagcctatgt ggtcctgagg ctgccctacc agctggccca gctgctctac     780

ctggccttcc tgtgggacgt ctactctggc tacctgctct gggaggccct ggtctactcc     840

gactacctga tcctactcaa cagctgcctc agccccttcc tctgcctcat ggccagtgcc     900

gacctccgga ccctgctgcg ctccgtgctc tcgtccttcg cggcagctct ctgcgaggag     960

cggccgggca gcttcacgcc cactgagcca cagacccagc tagattctga gggtccaact    1020

ctgccagagc cgatggcaga ggcccagtca cagatggatc ctgtggccca gcctcaggtg    1080

aaccccacac tccagccacg atcggatccc acagctcagc cacagctgaa ccctacggcc    1140

cagccacagt cggatcccac agcccagcca cagctgaacc tcatggccca gccacagtca    1200

gactctgtgg cccagccaca ggcagacact aacgtccaga cccctgcacc tgctgccagt    1260

tctgtgccca gtccctgtga tgaagcttcc ccaaccccat cctcgcatcc taccccaggg    1320

gcccctgagg acccagccac acctcctgcc tctgaaggag aaagccccag cagcaccccg    1380
```

ccagaggcgg ccccgggcgc aggccccacg tga 1413

<210> 94
<211> 419
<212> PRT
<213> H.Sapiens

<400> 94

Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro
1 5 10 15

Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp
20 25 30

Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro Ala Asn Gly
35 40 45

Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr
50 55 60

Arg Leu Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe
65 70 75 80

Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His
85 90 95

Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly
100 105 110

Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp
115 120 125

Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro
130 135 140

Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr
145 150 155 160

Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp
165 170 175

Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser
180 185 190

Leu Arg Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe Leu Leu Leu
195 200 205

Leu Val Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg Thr Cys His
210 215 220

Arg Gln Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg
225 230 235 240

Thr Ile Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr Gln Leu Ala
245 250 255

Gln Leu Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu
260 265 270

Leu Trp Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser
275 280 285

Cys Leu Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr
290 295 300

Leu Leu Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Glu Glu
305 310 315 320

Arg Pro Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser
325 330 335

Glu Gly Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met
340 345 350

Asp Pro Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln Pro Arg Ser
355 360 365

Asp Pro Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser
370 375 380

Asp Pro Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser
385 390 395 400

Asp Ser Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala
405 410 415

Pro Ala Ala

<210> 95
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 95
ttcaaagctt atggaatcat ctttctcatt tggagtgatc cttgctgtc 49

<210> 96
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 96
ttcactcgag ttagccatca aactctgagc tggagatagt gacgatgtg 49

<210> 97
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 97
gttcaaccga ctcatctatg cc 22

<210> 98
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 98
aaacttctct gcccttaccg tc 22

<210> 99
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 99
aaagcagcac cccgaatacc 20

<210> 100
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 100
catgatcaac ctgagcgtca c 21

<210> 101

<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 101
ttcaaagctt atggagtcgg ggctgctg 28

<210> 102
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 102
ttcactcgag tcagtctgca gccggttctg 30

<210> 103
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 103
gcatcctggc cgctatctgt gcactctacg 30

<210> 104
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 104
cgtagagtgc acagatagcg gccaggatgc 30

<210> 105
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 105
aaccccatca tctacacgc 19

<210> 106
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 106
tgcctgtgca gccgctgg 18

<210> 107
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 107
gcataagctt ccatgtacaa cgggtcgtgc tgc 33

<210> 108
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 108
gaattctaga tcagtgccac tcaacaatgt ggg 33

<210> 109
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature

<223> Novel Sequence

<400> 109
gaagcccagc actgtttacc 20

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 110
tgaaatacct gtccgcagcc 20

<210> 111
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 111
gatcaagctt atgacaggtg acttcccaag tatgc 35

<210> 112
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 112
gatcctcgag gctaacggca caaaacacaa ttcc 34

<210> 113
<211> [illegible]
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 113
cagcccaaac atccaagtc 19

<210> 114
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 114
accccactta atcagcctc 19

<210> 115
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 115
gatcgaattc gcaggagcaa tgaaaatcag gaac 34

<210> 116
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 116
gatcgaattc ttatatatgt tcagaaaaca aattcatgg 39

<210> 117
<211> 20
<212> DNA
<213> Artificial Sequence

<400> 117
tcagccccaa agccaaacac 20

<210> 118
<211> 22
<212> DNA
<213> Artificial Sequence

<400> 118
tcgcaggagc aatgaaaatc ag 22

<210> 119
<211> 19
<212> DNA
<213> Artificial Sequence

<400> 119
ctgaaagttg tcgctgacc 19

<210> 120
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 120
ggattatcca cactttgacc c 21

<210> 121
<211> 25
<212> DNA
<213> Artificial Sequence

<400> 121
aaataccatg aatgagccac tagac 25

<210> 122
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 122
gaatctcgag tcaagggttg tttgagtaac 30

<210> 123
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 123
ctgtctctct gtcctcttcc 20

<210> 124
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 124
gcaccgatct tcattgaatt tc 22

<210> 125
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 125
acttcaaaca acttcatacc cc 22

<210> 126
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 126
acacacagca tagtagcg 18

<210> 127
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 127

cagagcttga tgatgaggac 20

<210> 128
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 128
cccataggca gtagtagaag 20

<210> 129
<211> 9
<212> PRT
<213> Synthetic substrate peptide

<220>
<221> misc_feature
<223> Novel Sequence

<400> 129

Ala Pro Arg Thr Pro Gly Gly Arg Arg
1 5

<210> 130
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 130
gagtaatacg actcactata gggagaccgc gtgtctgcta gactctattt cc 52

<210> 131
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 131
tgctacactg atgcaactcc 20

<210> 132
<211> 48
<212> DNA
<213> Artificial Sequence

<400> 132
gcgtaatacg actcactata gggagacctg ccacactgat gcaactcc 48

<210> 133
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 133
gcgtgtctgc tagactctat ttcc 24

<210> 134
<211> 50
<212> DNA
<213> Artificial Sequence

<400> 134
gcgtaatacg actcactata gggagaccgc acgccactct ttactatccc 50

<210> 135
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 135
gcacaaaaca caattccata agcc 24

<210> 136
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 136
gcgtaatacg actcactata gggagaccgc acaaaacaca attccataag cc 52

<210> 137
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 137
gctacgccac tctttactat ccc 23

<210> 138
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 138
gcgtaatacg actcactata gggagacctt atgagcagca attcatccc 49

<210> 139
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 139
cacacccacc aagaaatcag 20

<210> 140
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 140
gcgtaatacg actcactata gggagaccca cacccaccaa gaaatcag 48

<210> 141

<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 141
ttatgagcag caattcatcc c 21

<210> 142
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 142
gcgtaatacg actcactata gggagacccg attatccaca ctttgaccc 49

<210> 143
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 143
ctgaaagttg tcgctgacc 19

<210> 144
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 144
gcgtaatacg actcactata gggagaccct gctgaaagtt gtcgctgacc 50

<210> 145
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 145
ggattatcca cactttgacc c 21

<210> 146
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 146
gcgtaatacg actcactata gggagaccct gtaaaattca cacaagcacc 50

<210> 147
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 147
agaagacaga gcaacctcc 19

<210> 148
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 148
ggcgtaatac gactcactat agggagacca gaagacagag caacctcc 48

<210> 149
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature

<223> Novel Sequence

<400> 149
ctgtaaaatt cacacaagca cc 22

<210> 150
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 150
gcatggatcc tctttgctgt atttcaccct c 31

<210> 151
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 151
gcatgaattc acaatgccag tgataaggaa g 31

<210> 152
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 152
gatcaagctt ggaatgatgc ccttttgcca c 31

<210> 153
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 153
gatcctcgag catcattcaa agtaggtgg 29

<210> 154
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 154
gatcctcgag ctatgaactc aattccaaaa ataatttaca cc 42

<210> 155
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 155
gctacttgaa ctctacattt aatccaatgg tttatgcatt tttctatcc 49

<210> 156
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 156
ggatagaaaa atgcataaac cattggatta aatgtagagt tcaagtagc 49

<210> 157
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 157
gatcgaattc atggacacta ccatggaagc tgacc 35

<210> 158
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 158
gatcctcgag tcacgtgggg cctgcgcccg g 31

<210> 159
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 159
gcgtaatacg actcactata gggagaccgc gtgtctgcta gactctattt cc 52

<210> 160
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 160
tgccacactg atgcaactcc 20

<210> 161
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 161
gcgtaatacg actcactata gggagacctg ccacactgat gcaactcc 48

<210> 162
<211> 24

<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 162
gcgtgtctgc tagactctat ttcc 24

<210> 163
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 163
gcgtaatacg actcactata gggagaccgc acgccactct ttactatccc 50

<210> 164
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 164
gcacaaaaca caattccata agcc 24

<210> 165
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 165
gcgtaatacg actcactata gggagaccgc acaaaacaca attccataag cc 52

<210> 166
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 166
gctacgccac tctttactat ccc 23

<210> 167
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 167
gcgtaatacg actcactata gggagacctt atgagcagca attcatccc 49

<210> 168
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 168
cacacccacc aagaaatcag 20

<210> 169
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 169
gcgtaatacg actcactata gggagaccca cacccaccaa gaaatcag 48

<210> 170
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 170
ttatgagcag caattcatcc c 21

<210> 171
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 171
gcgtaatacg actcactata gggagacccg attatccaca ctttgaccc 49

<210> 172
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 172
ctgaaagttg tcgctgacc 19

<210> 173
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 173
gcgtaatacg actcactata gggagacccct gctgaaagtt gtcgctgacc 50

<210> 174
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 174

cgattatcca cactttgacc c 21

<210> 175
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 175
gcgtaatacg actcactata gggagaccct gtaaaattca cacaagcacc 50

<210> 176
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 176
agaagacaga gcaacctcc 19

<210> 177
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 177
gcgtaatacg actcactata gggagaccag aagacagagc aacctcc 47

<210> 178
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 178
ctgtaaaatt cacacaagca cc 22

<210> 179
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 179
gcatggatcc tctttgctgt atttcaccct c 31

<210> 180
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 180
gcatgaattc acaatgccag tgataaggaa g 31

<210> 181
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 181
tcagccccaa agccaaacac 20

<210> 182
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 182
ccgtaggagc aatgaaaatc ag 22

<210> 183
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 183
ctgtctctct gtcctcttcc 20

<210> 184
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 184
gtacagatct tcattgaatt tc 22

<210> 185
<211> 1188
<212> DNA
<213> H.Sapiens

<400> 185
aggctcgcgc ccgaagcaga gccatgagaa ccccagggtg cctggcgagc cgctagcgcc 60
atgggccccg gcgaggcgct gctggcgggt ctcctggtga tggtactggc cgtggcgctg 120
ctatccaacg cactggtgct gctttgttgc gcctacagcg ctgagctccg cactcgagcc 180
tcaggcgtcc tcctggtgaa tctgtctctg ggccacctgc tgctggcggc gctggacatg 240
cccttcacgc tgctcggtgt gatgcgcggg cggacaccgt cggcgcccgg cgcatgccaa 300
gtcattggct tcctggacac cttcctggcg tccaacgcgg cgctgagcgt ggcggcgctg 360
agcgcagacc agtggctggc agtgggcttc ccactgcgct acgccggacg cctgcgaccg 420
cgctatgccg gcctgctgct gggctgtgcc tggggacagt cgctggcctt ctcaggcgct 480
gcacttggct gctcgtggct tggctacagc agcgccttcg cgtcctgttc gctgcgcctg 540
ccgcccgagc ctgagcgtcc gcgcttcgca gccttcaccg ccacgctcca tgccgtgggc 600
ttcgtgctgc cgctggcggt gctctgcctc acctcgctcc aggtgcaccg ggtggcacgc 660
agacactgcc agcgcatgga caccgtcacc atgaaggcgc tcgcgctgct cgccgacctg 720
caccccagtg tgcggcagcg ctgcctcatc cagcagaagc ggcgccgcca ccgcgccacc 780
aggaagattg gcattgctat tgcgaccttc ctcatctgct ttgccccgta tgtcatgacc 840

aggctggcgg agctcgtgcc cttcgtcacc gtgaacgccc agtggggcat cctcagcaag 900

tgcctgacct acagcaaggc ggtggccgac ccgttcacgt actctctgct ccgccggccg 960

ttccgccaag tcctggccgg catggtgcac cggctgctga agagaacccc gcgcccagca 1020

tccacccatg acagctctct ggatgtggcc ggcatggtgc accagctgct gaagagaacc 1080

ccgcgcccag cgtccaccca caacggctct gtggacacag agaatgattc ctgcctgcag 1140

cagacacact gagggcctgg cagggctcat cgcccccacc ttctaaga 1188

<210> 186
<211> 363
<212> PRT
<213> H.Sapiens

<400> 186

Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met Val Leu
1 5 10 15

Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys Ala Tyr
20 25 30

Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val Asn Leu
35 40 45

Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe Thr Leu
50 55 60

Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala Cys Gln
65 70 75 80

Val Ile Gly Phe Leu Asp Thr Phe Leu Ala Ser Asn Ala Ala Leu Ser
85 90 95

Val Ala Ala Leu Ser Ala Asp Gln Trp Leu Ala Val Gly Phe Pro Leu
100 105 110

Arg Tyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu Leu Gly
115 120 125

Cys Ala Trp Gly Gln Ser Leu Ala Phe Ser Gly Ala Ala Leu Gly Cys
130 135 140

Ser Trp Leu Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu Arg Leu
145 150 155 160

Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Phe Thr Ala Thr Leu
165 170 175

His Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cys Leu Thr Ser
180 185 190

Leu Gln Val His Arg Val Ala Arg Arg His Cys Gln Arg Met Asp Thr
195 200 205

Val Thr Met Lys Ala Leu Ala Leu Leu Ala Asp Leu His Pro Ser Val
    210   215   220

Arg Gln Arg Cys Leu Ile Gln Gln Lys Arg Arg Arg His Arg Ala Thr
225   230   235   240

Arg Lys Ile Gly Ile Ala Ile Ala Thr Phe Leu Ile Cys Phe Ala Pro
                245   250   255

Tyr Val Met Thr Arg Leu Ala Glu Leu Val Pro Phe Val Thr Val Asn
            260   265   270

Ala Gln Trp Gly Ile Leu Ser Lys Cys Leu Thr Tyr Ser Lys Ala Val
        275   280   285

Ala Asp Pro Phe Thr Tyr Ser Leu Leu Arg Arg Pro Phe Arg Gln Val
    290   295   300

Leu Ala Gly Met Val His Arg Leu Leu Lys Arg Thr Pro Arg Pro Ala
305   310   315   320

Ser Thr His Asp Ser Ser Leu Asp Val Ala Gly Met Val His Gln Leu
                325   330   335

Leu Lys Arg Thr Pro Arg Pro Ala Ser Thr His Asn Gly Ser Val Asp
            340   345   350

Thr Glu Asn Asp Ser Cys Leu Gln Gln Thr His
        355   360

<210> 187
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 187
gcataagctt gccatgggcc ccggcgagg 29

<210> 188
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 188
gcattctaga cctcagtgtg tctgctgc 28

<210> 189

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 189
tgctgctttg ttgcgcctac 20

<210> 190
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 190
ttggacgcca ggaaggtg 18

SEQUENCE LISTING

<110> Pharmacia & Upjohn Company
Vogeli, Gabriel
Huff, Rita
Sejlitz, Torsten
Lind, Peter
Slightom, Jerry
Schellin, Kathleen
Bannigan, Chris
Buff, Valerie
Kaytes, Paul
Wood, Linda
Parodi, Luis
Hiebsch, Ronald

<120> Novel G Protein Coupled Receptors

<130> 04SP1PHRM296

<150> 60/165,830
<151> 1999-11-16

<150> 60/198,568
<151> 2000-04-20

<150> 60/166,071
<151> 1999-11-17

<150> 60/166,676
<151> 1999-11-19

<150> 60/173,396
<151> 1999-12-28

<150> 60/184,129
<151> 2000-02-22

<150> 60/185,421
<151> 2000-02-28

<150> 60/185,554
<151> 2000-02-28

<150> 60/186,530
<151> 2000-03-02

<150> 60/186,811
<151> 2000-03-03

<150> 60/188,114
<151> 2000-03-09

<150> 60/190,310
<151> 2000-03-17

<150> 60/190,800
<151> 2000-03-21

<150> 60/201,190
<151> 2000-05-02

<150> 60/203,111
<151> 2000-05-08

<150> 60/207,094
<151> 2000-05-25

<160> 190

<170> PatentIn version 3.0

<210> 1
<211> 1182
<212> DNA
<213> H.Sapiens

<400> 1
```
gtctgggggt ggggatgct gggacagggg tcaattgcct gaagcaagtg ctctcatccc      60

cctagctcct gctgatctag ttgggctcc agagtgggga gggagaaggc acttttgaaac     120

ttctctgccc ttaccgtctt agcccatcaaa ctctgagctg gagatagtga cgatgtgaca    180

ggaactttcc ctgggcctct ctgggccaca attcctggcc gagagaaaga ggaggaatga     240

ggtgagcacc ttcttcactc ctagggccat gtggtagagc tgcagtcgca cctccttctg     300

ccaataggca tagatgagtg ggttgagcag ggagttgccc acgccgagca gccacaggta     360

ccgttccagc actaggtaga ggtgacccto ctgggaggcc acctgccaae tgccagtgat     420

aaggaaggcg gtccaggata gagcaagct cccaatgaga acagacacag tacggagagc      480

tttgaagtcg ctgggagtcc gtggggatcg ataacctcca gccatggctc ctgcatgttc     540

catctttccs atctgctggc tgtgcatgga ggcaatcttg agcatgtcgc agtagaagaa     600

gacaaagagg agcatggctg ggaagaagcc aacgcaggag agggtcagca cgaagtgagg     660

gtgaaataca gcaaagaagc tgcactgccc ttgtaggca gtctgctgge acatgggcct      720

tccgagtggg aggaagccca tgaggtaaga cactaacac agcccggcaa tgcaggcccc      780

ggccacgaac ccacccatga tcttcaagta gcggaagggc tgcttgatgg caaggtacct     840

gtcaaaggtg atcagcatga ccgtgaggac agaggcagct gcggaggaag tgacaaatgc     900

catccgcagg ctgcacaggg tcttctgtgt gggccgagaa gggctggaga gctggtctgt     960

gagtaggcca gagatggcca caccaatcac ggtgtcagcc acagcccgat tcaagtgaa     1020

gcagagactg acaccatcat tcttgtggat caacagcagc acagcccacag ccactagtgt    1080

gttagtagca atgatgaggg aggccaggac agcaaggatc actccaaatg agaaagatga    1140

ttccatgtct cgaagtggca ggacttcact tacccaggca tg                        1182
```

<210> 2
<211> 335
<212> PRT
<213> H.Sapiens

<400> 2

```
Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser
1               5                   10                  15

Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu
            20                  25                  30

Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala
        35                  40                  45

Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp
    50                  55                  60

Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu
65                  70                  75                  80

Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val
                85                  90                  95

Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg
            100                 105                 110

Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly
        115                 120                 125

Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro
    130                 135                 140

Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val
145                 150                 155                 160

Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro
                165                 170                 175

Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala
            180                 185                 190

Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met
        195                 200                 205

Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu
    210                 215                 220

Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro
225                 230                 235                 240

Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu
                245                 250                 255

Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser
            260                 265                 270
```

Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu
275 280 285

Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe
290 295 300

Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu
305 310 315 320

Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly
325 330 335

<210> 3
<211> 657
<212> DNA
<213> H.Sapiens

<400> 3
cagcgcgagc gccttcatgg tgacggtgtc catgcgctgg cagtgtctgc gtgccacccg 60

gtgcacctgg agcgaggtga ggcagagcac cgccagcggc agcacgaagc ccacggcatg 120

gagcgtgacg ctcaaggctg cgaagcgcgg acgctcaggc tcgggcggca ggcgcagcga 180

[illegible] aaggcgctgc tgtagccaag ccacgagcag ccaagtgcag cgcctcagaa 240

ggccagcgac tgtccccagc cacagcccag cagcaggccg gcatagcccg gtcgcaggcg 300

tccggcgtag cgcagtgggc agcccactgc cagccactgg tctgcgctca gcgcccccac 360

gctcagcgcc gccttgcacg ccaggaaggt gtccaggaag ccaatgactt ggcatgcgcc 420

gggcgccgac ggtgtccgcc cgcgcatcac accgagcagc gtgaagggca tgtccagcgc 480

cgccagcagc aggtggccca gagacagatt cccaaggagg cgccctgagg ctcgagtgcg 540

gagctcagcg ctgtaggcgc aaccaagcag caccagtgcg ttgcatagca gcgccacggc 600

cagtaccatc accaggagac ccgccagcag cgcctcgccg gggcccatgg cgctagc 657

<210> 4
<211> 217
<212> PRT
<213> H.Sapiens

<400> 4

Ser Ala Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met
1 5 10 15

Val Leu Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys
20 25 30

Ala Tyr Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val
35 40 45

Asn Leu Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe
50 55 60

Thr Leu Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala
65 70 75 80

Cys Gln Val Ile Gly Phe Leu Asp Thr Phe Leu Ala Ser Asn Ala Ala
85 90 95

Leu Ser Val Ala Ala Leu Ser Ala Asp Gln Trp Leu Ala Val Gly Phe
100 105 110

Pro Leu Arg Tyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu
115 120 125

Leu Gly Cys Ala Trp Gly Gln Ser Leu Ala Phe Ser Gly Ala Ala Leu
130 135 140

Gly Cys Ser Trp Leu Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu
145 150 155 160

Arg Leu Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Phe Thr Ala
165 170 175

Thr Leu His Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cys Leu
180 185 190

Thr Ser Leu Gln Val His Arg Val Ala Arg Arg His Cys Gln Arg Met
195 200 205

Asp Thr Val Thr Met Lys Ala Leu Ala
210 215

<210> 5
<211> 222
<212> DNA
<213> H.Sapiens

<400> 5
tgtgcaggtg tgatctccat tcctttgtac atccctcaca cgctgttcga atgggatttt 60

ggaaaggaaa tctgtgtatt ttggctcact actgactatc tgttatgtac agcatctgta 120

tataacattg tcctcatcag ctatgatcga tacctgtcag tctcaaatgc tgtaagtcga 180

acacattaat ttatccccct tagaagatta tgtaaatgta ta 222

<210> 6
<211> 73
<212> PRT
<213> H.Sapiens

<400> 6

Cys Ala Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe
1 5 10 15

Glu Trp Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp
20 25 30

Tyr Leu Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr
35 40 45

Asp Arg Tyr Leu Ser Val Ser Asn Ala Val Ser Arg Thr His Phe Ile
50 55 60

Pro Leu Arg Arg Leu Cys Lys Cys Ile
65 70

<210> 7
<211> 507
<212> DNA
<213> H.Sapiens

<400> 7
gacgtcgaag caggtgatga tgcccagggc gtgcaccggg taggtgagat cggtgcggcg 60

cagcggggac agggcggtca ggagcagcag ccaggtccct gcacacgcgg ccaccgcgta 120

acgacggcgg cgccagcgct tggagctgag cgggtacagg atccccagga agcgctccac 180

gctgatacag gtcatggtga ggatgctgga atacatgttt gcgtaaaacg ccacggtcac 240

cacgttgcaa agcagcaccc cgaataccca gtggtggcgg ttgcaatgct agtagatttg 300

gaaaggtaac atgctggcca gcatcaggtc cctgacgctc aggttgatcc tgaagatgac 360

cgacggggat ctgggcccca tgcgccggca cagcacccac agagatgaga ggttgcccgg 420

gatgctgacc gccgccacca gcgagtacac cacgggcagg gccaccgcga tcgccgggtt 480

ccgcagcatc tgcagcgtcg cgttgtc 507

<210> 8
<211> 169
<212> PRT
<213> H.Sapiens

<400> 8

Asp Asn Ala Thr Leu Gln Met Leu Arg Asn Pro Ala Ile Ala Val Ala
1 5 10 15

Leu Pro Val Val Tyr Ser Leu Val Ala Ala Val Ser Ile Pro Gly Asn
20 25 30

Leu Phe Ser Leu Trp Val Leu Cys Arg Arg Met Gly Pro Arg Ser Pro
35 40 45

Ser Val Ile Phe Met Ile Asn Leu Ser Val Thr Asp Leu Met Leu Ala
50 55 60

Ser Val Leu Pro Phe Gln Ile Tyr Tyr His Cys Asn Arg His His Trp
65 70 75 80

Val Phe Gly Val Leu Cys Asn Leu Val Val Thr Val Ala Phe Tyr Ala
85 90 95

```
Asn Met Tyr Ser Ser Ile Leu Thr Met Thr Cys Ile Ser Val Glu Arg
            100                 105                 110

Phe Leu Gly Ile Leu Tyr Pro Leu Ser Ser Lys Arg Trp Arg Arg Arg
        115                 120                 125

Arg Tyr Ala Val Ala Ala Cys Ala Gly Thr Trp Leu Leu Leu Leu Thr
    130                 135                 140

Ala Leu Ser Pro Leu Ala Arg Thr Asp Leu Thr Tyr Pro Val His Ala
145                 150                 155                 160

Leu Gly Ile Ile Thr Cys Phe Asp Val
                165
```

<210> 9
<211> 270
<212> DNA
<213> H.Sapiens

<400> 9

```
cccatgttcc tgctcctggg cagcctcacg ttgtcggatc tgctggcagg cgccgcctac      60

gccgccaaca tcctactgtc ggggccgctc acgctgaaac tgtcccccgc gctctggttc     120

gcacgggagg gaggcgtctt cgtggcactc actgcgtccg tgctgagcct cctgggcctc     180

gcgctggagc gcagcctcac catggcgcgc aggggcccgg cgcccgtctc cagtcggggg     240

cgcacgctgg cgatggcagc cgcggcctgg                                      270
```

<210> 10
<211> 90
<212> PRT
<213> H.Sapiens

<400> 10

```
Pro Met Phe Leu Leu Leu Gly Ser Leu Thr Leu Ser Asp Leu Leu Ala
1               5                   10                  15

Gly Ala Ala Tyr Ala Ala Asn Ile Leu Leu Ser Gly Pro Leu Thr Leu
            20                  25                  30

Lys Leu Ser Pro Ala Leu Trp Phe Ala Arg Glu Gly Gly Val Phe Val
        35                  40                  45

Ala Leu Thr Ala Ser Val Leu Ser Leu Leu Gly Ile Ala Leu Glu Arg
    50                  55                  60

Ser Leu Thr Met Ala Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Gly
65                  70                  75                  80

Arg Thr Leu Ala Met Ala Ala Ala Ala Trp
                85                  90
```

<210> 11
<211> 688

<212> DNA
<213> H.Sapiens

<400> 11
ctgctcattg tggcctttgt gctgggcgca ctaggcaatg gggtcggcct gtgtggtttc 60

tgcttccaca tgaagacctg gaagcccagc actgtttacc ttttcaattt ggccgtggct 120

gatttcctcc ttatgatctg cctgcctttt cggacagact attacctcag acgtagacac 180

tgggcttttg gggacattcc ctgccgagtg gggctcttca cgttggccat gaacagggcc 240

gggagcatcg tgttccttac ggtggtggct gcggacaggt atttcaaggt ggtccacccc 300

caccacgcgg tgaacactat ctccacccgg gtggcggctg gcatcgtctg caccctgtgg 360

gccctggtcc tcctgggaac agtgtatctt ttgctggaga accatctctg cgtgcaagag 420

acggccgtct cctgtgagag cttcatcatg gagtcggcca atggctggca tgacatcatg 480

ttccagctgg acttctttat gcccctcggc atcatcttat tttgctcctt caagattgtt 540

tggagcctga ggcggaggca gcagctggcc agacaggctc ggatgaagaa ggcgacccgg 600

ttcatcatgg tggtggcaat tgtgttcatc acatgctacc tgcccagcgt gtctgctaga 660

ctctatttcc tctggacggt gccctcgagt gcctgcgatc cctctgtcca tggggccctg 720

cacataaccc tcagcttcac ctacatgaac agcatgctgg atcccctggt gtattatttt 780

tcaagcccct ccttccccaa attctacaac aagctcaaaa tctgcagtct gaaacccaag 840

cagccaggac actcaaaaac acaaaggccg gaagagatgc caatttcg 888

<210> 12
<211> 296
<212> PRT
<213> H.Sapiens

<400> 12

Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly Asn Gly Val Ala
1 5 10 15

Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys Pro Ser Thr Val
20 25 30

Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu Met Ile Cys Leu
35 40 45

Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His Trp Ala Phe Gly
50 55 60

Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala Met Asn Arg Ala
65 70 75 80

Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp Arg Tyr Phe Lys
85 90 95

Val Val His Pro His His Ala Val Asn Thr Ile Ser Thr Arg Val Ala
100 105 110

Ala Gly Ile Val Cys Thr Leu Trp Ala Leu Val Ile Leu Gly Thr Val
115 120 125

Tyr Leu Leu Leu Glu Asn His Leu Cys Val Gln Glu Thr Ala Val Ser
130 135 140

Cys Glu Ser Phe Ile Met Glu Ser Ala Asn Gly Trp His Asp Ile Met
145 150 155 160

Phe Gln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile Leu Phe Cys Ser
165 170 175

Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln Leu Ala Arg Gln
180 185 190

Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val Val Ala Ile Val
195 200 205

Phe Ile Thr Cys Tyr Leu Pro Ser Val Ser Ala Arg Leu Tyr Phe Leu
210 215 220

Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val His Gly Ala Leu
225 230 235 240

His Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met Leu Asp Pro Leu
245 250 255

Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe Tyr Asn Lys Leu
260 265 270

Lys Ile Cys Ser Leu Lys Pro Lys Gln Pro Gly His Ser Lys Thr Gln
275 280 285

Arg Pro Glu Glu Met Pro Ile Ser
290 295

<210> 13
<211> 510
<212> DNA
<213> H.Sapiens

<400> 13
tggagctgtg ccacccacta tctggtgaac ctgatggtgg ccgacctgct ttatgtgcta 60

ttgcccttcc tcatcatcac ctactcacta gatgacaggt ggcccttcgg ggagctgctc 120

tgcaagctgg tgcacttcct gttctatatc aacctttacg gcagcatcct gctgctgacc 180

tgcatctctg tgcaccagtt cctaggtgtg tgccaccgac tgtgttcgct gccctaccgg 240

acccgcaggc atgcctggct gggcaccagc accacctggg ccctggtggt cctccagctg 300

ctgcccacac tggccttctc ccacacggac tacatcaatg gccagatgat ctggtatgac 360

atgaccagcc aagagaattt tgatcggctt tttgcctacg gcatagttct gacattgtct 420

ggctttcttt ccctccttgg tcattttggt gtgctattca ctgatggtca ggagcctgat 480

caagccagag gagaacctca tgaggacagg 510

<210> 14
<211> 170
<212> PRT
<213> H.Sapiens

<400> 14

Trp Ser Cys Ala Thr Thr Tyr Leu Val Asn Leu Met Val Ala Asp Leu
1 5 10 15

Leu Tyr Val Leu Leu Pro Phe Leu Ile Ile Thr Tyr Ser Leu Asp Asp
20 25 30

Arg Trp Pro Phe Gly Glu Leu Leu Cys Lys Leu Val His Phe Leu Phe
35 40 45

Tyr Ile Asn Leu Tyr Gly Ser Ile Leu Leu Leu Thr Cys Ile Ser Val
50 55 60

His Gln Phe Leu Gly Val Cys His Pro Leu Cys Ser Leu Pro Tyr Arg
65 70 75 80

Thr Arg Arg His Ala Trp Leu Gly Thr Ser Thr Thr Trp Ala Leu Val
85 90 95

Val Leu Gln Leu Leu Pro Thr Leu Ala Phe Ser His Thr Asp Tyr Ile
100 105 110

Asn Gly Gln Met Ile Trp Tyr Asp Met Thr Ser Gln Glu Asn Phe Asp
115 120 125

Arg Leu Phe Ala Tyr Gly Ile Val Leu Thr Leu Ser Gly Phe Leu Ser
130 135 140

Leu Leu Gly His Phe Gly Val Leu Phe Thr Asp Gly Gln Glu Pro Asp
145 150 155 160

Gln Ala Arg Gly Glu Pro His Glu Asp Arg
165 170

<210> 15
<211> 894
<212> DNA
<213> H.Sapiens

<220>
<221> misc_feature
<222> (431)..(461)
<223> n is any nucleotide

<400> 15
ccaccacgcg cagcacgccg acagggcctc tccctcccat tctcccgcag gccccgacga 60

ccacgctgcc tccagccggt cggcaaacta gggcagctcg cagcccacga acagcagccc 120

cagcagctgg ctcatcttca ggctctgcac cttggcgcgg ggcatcgcgc tgggcgcacg 180

ggctccacct gggctcgccg accaggccgc tgcacccgct ggggccttca gccggtgccg 240

ccaccagccg gagagtcggt ggccacaagc gacacccatc atcttaacag gcgcgacgaa 300

gcccgcgacg gcctcataga acgcgtacac ctgcacgtgc cagcgctgca gcagcgcgaa 360

gatccagtgg cagcgacgca tccccggcca ggctcgggcg gagagtggcg cgcctggctc 420

cagagacgtt nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nagtactagc gcaccacaaa 480

cccggacccc cgcgccagca gcagtgcccg ccgccagccc aggcggcgga gggcacgcgc 540

gggcagcggc cggccgtgcc gaagacgcac ccggcgccgg cgctcgaggg cgatgagcac 600

cacgaggtgg gccgaggcgc cccgcccgga tgcctgcagc agctgcagga agcggccgcc 660

caggtccccc gtggccgcgc ggggctcgcc cagcagttcc caggccagct gtgacagcgc 720

cgtgcccccg cacgcgtacc ggtccgcccg gcccagctgc accagcagga agtccatctt 780

gcgacgcttn nnnnnnnnnn nnnnnnnnnn nnnnnnnnac aggcggcacg gcactgtggt 840

gttgcctgcc accgccacca ccaggatgac ccccaggaac accaggcgga cgcg 894

<210> 16
<211> 296
<212> PRT
<213> H.Sapiens

<220>
<221> UNSURE
<222> (26)..(35)
<223> Xaa is unknown

<220>
<221> UNSURE
<222> (144)..(154)
<223> Xaa is Unknown

<400> 16

Arg Val Arg Leu Val Phe Leu Gly Val Ile Leu Val Val Ala Val Ala
1 5 10 15

Gly Asn Thr Thr Val Leu Cys Arg Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa
20 25 30

Xaa Xaa Xaa Lys Arg Arg Lys Met Asp Phe Leu Leu Val Gln Leu Ala
35 40 45

Leu Ala Asp Leu Tyr Ala Cys Gly Gly Thr Ala Leu Ser Gln Leu Ala

50 55 60

Trp Glu Leu Leu Gly Glu Pro Arg Ala Ala Thr Gly Asp Leu Ala Cys
65 70 75 80

Arg Phe Leu Gln Leu Leu Gln Ala Ser Gly Arg Gly Ala Ser Ala His
85 90 95

Leu Val Val Leu Ile Ala Leu Glu Arg Arg Arg Ala Val Arg Leu Pro
100 105 110

His Gly Arg Pro Leu Pro Ala Arg Ala Leu Ala Ala Leu Gly Trp Leu
115 120 125

Leu Ala Leu Leu Leu Ala Arg Gly Ser Gly Phe Val Val Arg Tyr Xaa
130 135 140

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Thr Ser Leu Gln Pro Gly
145 150 155 160

Ala Pro Leu Ser Ala Arg Ala Trp Pro Gly Met Arg Arg Cys His Trp
165 170 175

Ile Phe Ala Leu Leu Gln Arg Trp His Val Gln Val Tyr Ala Phe Tyr
180 185 190

Glu Ala Val Ala Gly Phe Val Ala Pro Val Lys Ile Met Gly Val Ala
195 200 205

Cys Gly His Leu Leu Ser Val Trp Trp Arg His Arg Leu Lys Ala Pro
210 215 220

Ala Gly Ala Ala Ala Trp Ser Ala Ser Pro Gly Gly Ala Arg Ala Pro
225 230 235 240

Ser Ala Met Pro Arg Ala Lys Val Gln Ser Leu Lys Met Ser Gln Leu
245 250 255

Leu Gly Leu Leu Phe Val Gly Cys Glu Leu Pro Phe Ala Asp Arg Leu
260 265 270

Glu Ala Ala Trp Ser Ser Gly Pro Ala Gly Glu Trp Glu Gly Glu Ala
275 280 285

Leu Ser Ala Cys Cys Ala Trp Trp
290 295

<210> 17
<211> 801
<212> DNA
<213> H.Sapiens

<400> 17
tctaagtttt tctctgaact ttgagcctgt gaaaaagaa gggatgctgc ctcaggccac 60

cccagcctag atactcactc tgagtgccat gaggtagtag aggacactga tgacagtcat 120

ggggaggagg tagaatagga aggaggtgac ctgggtgatg aaattgtaga tccacatggg 180

cctgatgacc ctacaggtgg ccgaacctgg gacccgggac ccattgggga agtagtggaa 240

cttgatgcca tggatgctgg tgttgggcag ggagaagagc acgggagagc cccagacgat 300

gccgaggatc ctgagggccc ggcgccgggt gctctgcagt ttggcgcgga acgggtgtag 360

gatggccacg tagcgctcca cgctgacggt ggtgatgctg aggatggagg cgaagcccac 420

ggtctcaaag agggccgtct tgaagtagca gcccacgggc ccgaacaaga aagggtagtt 480

gcgccacatc tcatagacct ccagggccat tccaaggagc aggaccagga ggtcagagac 540

cgccaggctg aagaggtagt agttggtggg cgtcttcata gcctggtgct gcagaatcac 600

caggcagacc aggacattgc caatgacccc caccacaaaa attggcacat acaccacaga 660

cacgggggag aagaagtggc tgcgccgagg tccgcagagg aaggccagat actcctcggt 720

gctgttcagg tgtttctgga atggatcttc tagtttctgc tggtagatcc aggaagcctt 780

ctgaagtttt tccatccctg a 801

<210> 18
<211> 249
<212> PRT
<213> H.Sapiens

<400> 18

Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln Lys
1 5 10 15

Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr Leu
20 25 30

Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val Ser
35 40 45

Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val Leu
50 55 60

Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Asn Thr
65 70 75 80

Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu Leu
85 90 95

Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe Leu
100 105 110

Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr Val
115 120 125

Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg Tyr
130 135 140

Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg Arg
145 150 155 160

```
Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu Phe
                165                 170                 175

Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe Pro
            180                 185                 190

Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys Pro
        195                 200                 205

Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe Tyr
    210                 215                 220

Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala Leu
225                 230                 235                 240

Arg Val Ser Ile Ala Gly Val Ala Gly
                245

<210>  19
<211>  222
<212>  DNA
<213>  H.Sapiens

<400>  19
atcaagatga ttttgctat cgtgcaaatt attggatttt ccaactccat ctgtaatccc      60

attgtctatg catttatgaa tgaaaacttc aaaaaaaatg ttttgtctgc agtttgttat     120

tgcatagtaa ataaaacctt ctctccagcc caaaggcatg gaaattcagg aattacaatg     180

atgaggaaga aagcaaagtt ttccctcaga gagaatccag tg                        222


<210>  20
<211>  73
<212>  PRT
<213>  H.Sapiens

<400>  20

Ile Lys Met Ile Phe Ala Ile Val Gln Ile Ile Gly Phe Ser Asn Ser
1               5                   10                  15

Ile Cys Asn Pro Ile Val Tyr Ala Phe Met Asn Glu Asn Phe Lys Lys
            20                  25                  30

Asn Val Leu Ser Ala Val Cys Tyr Cys Ile Val Asn Lys Thr Phe Ser
        35                  40                  45

Pro Ala Gln Arg His Gly Asn Ser Gly Ile Thr Met Met Arg Lys Lys
    50                  55                  60

Ala Lys Phe Ser Leu Arg Glu Asn Pro
65                  70

<210>  21
<211>  447
<212>  DNA
```

<213> H.Sapiens

<400> 21
gccacagcat gcagttttct gtagaattcc actttgtctt tgcacttgaa gaagctgagg 60

tatctggtga ccaggatcac cacatagaat aggaaccgtg aggtacatgt ggatgtgcag 120

catggcactc acaaatttgc agaagggcag cccaaacatc caagtcttct tgatgaggta 180

ggtcaagcga aatggcactg tcagcagaaa aacgctgtgg accaccacca agttaatgac 240

cgccatggtg gtcactgacc gggtgttcat tttcaccagg aggaaaagac tggaaatgac 300

acccaccagc ccgccaataa gcactatgaa gtagaggctg attaagtggg gtgtccctat 360

aggatcgcaa gaggaattcc tggaggtatt gtggccaggc atacttggga agtcacctgg 420

aggagaaaaa gcaccagagt aactgac 447

<210> 22
<211> 149
<212> PRT
<213> H.Sapiens

<400> 22

Val Ser Tyr Ser Gly Ala Phe Ser Pro Pro Gly Asp Phe Pro Ser Met
1 5 10 15

Pro Gly His Asn Thr Ser Arg Asn Ser Ser Cys Asp Pro Ile Val Thr
20 25 30

Pro His Leu Ile Ser Leu Tyr Phe Ile Val Leu Ile Gly Gly Leu Val
35 40 45

Gly Val Ile Ser Ile Leu Phe Leu Leu Val Lys Met Asn Thr Arg Ser
50 55 60

Val Thr Thr Met Ala Val Ile Asn Leu Val Val Val His Ser Val Phe
65 70 75 80

Leu Leu Thr Val Pro Phe Arg Leu Thr Tyr Leu Ile Lys Lys Thr Trp
85 90 95

Met Phe Gly Leu Pro Phe Cys Lys Phe Val Ser Ala Met Leu His Ile
100 105 110

His Met Tyr Leu Thr Val Pro Ile Leu Cys Gly Asp Pro Gly His Gln
115 120 125

Ile Pro His Leu Leu Gln Val Gln Arg Gln Ser Gly Ile Leu Gln Lys
130 135 140

Thr Ala Cys Cys Gly
145

<210> 23
<211> 222

<212> DNA
<213> H.Sapiens

<400> 23
actggccaag gtcagggcat cgactgaggc tagaaggcca caggaaatgc cagtcaaggt 60

gttggcgcct gcaatcgcac ctaccacaan cttgaccggg ggcagggggg caggccggcc 120

agcgaacacg gtcagcagcc ccagtccatt gcagagcacg gagagcaaca cgatggccca 180

cacggccacg cggatgcccc agctttcaaa gaggtactca ca 222

<210> 24
<211> 74
<212> PRT
<213> H.Sapiens

<400> 24

Cys Glu Tyr Leu Phe Glu Ser Trp Gly Ile Arg Leu Ala Val Trp Ala
1 5 10 15

Ile Val Leu Leu Ser Val Leu Cys Asn Gly Leu Val Leu Leu Thr Val
20 25 30

Phe Ala Gly Gly Pro Ala Pro Leu Pro Pro Val Lys Phe Val Val Gly
35 40 45

Ala Ile Ala Gly Ala Asn Thr Leu Thr Gly Ile Ser Cys Gly Leu Leu
50 55 60

Ala Ser Val Asp Ala Leu Thr Leu Val Ser
65 70

<210> 25
<211> 246
<212> DNA
<213> H.Sapiens

<400> 25
gacccgatcc tctacacgct cacaaaccgc gacctgcgcc accgcgtcct gcgcctggtc 60

tgctgcggac gccactcctg cggcagagac ccgagtggct cccagcagtc ggcgagcgcg 120

gctgaggctt ccgggggcct gcgccgctgc ctgcccccgg gccttgatgg gagcttcagc 180

ggctcggagc gctcatcgcc ccagcgcgcc gggctggaca ccagcggctc caccggcagc 240

cccggt 246

<210> 26
<211> 82
<212> PRT
<213> H.Sapiens

<400> 26

Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg His Ala Leu
1 5 10 15

Leu Arg Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg Asp Pro Ser
20 25 30

Gly Ser Gln Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly Gly Leu Arg
35 40 45

Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly Ser Glu Arg
50 55 60

Ser Ser Pro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser Thr Gly Ser
65 70 75 80

Pro Gly

<210> 27
<211> 420
<212> DNA
<213> H.Sapiens

<220>
<221> misc_feature
<222> (81)..(106)
<223> n is any nucleic acid

<400> 27
cgtgaagaac agcgccacca tgaccagcat gtgcaccacg cgcgctctgc gccgcgatgc 60

tcgcgggtcc gcagcctcct nnnnnnnnnn nnnnnnnnnn nnnnnntggc agagcttgcg 120

cgcgatgcgg gcgtacatga ccacgatgag cgccagcggc gccaggtaga tgtgcgagaa 180

gagcacagtg gtgtagaccc tgcgcatgcc cttctcgggc cagccctccc agcacgagta 240

gagagggtag gagcggttgc gggcgtccac catgaagtgg tgctcctcac gggtgacggt 300

cagcgtgacg gccgagggac acatgatgag cagcgccagg gcccagatga cggcgatggt 360

gacgagcgcc ttccgcaggg tcagcttctc gcggaaaggg tgcacgatgc agcggaacct 420

<210> 28
<211> 139
<212> PRT
<213> H.Sapiens

<220>
<221> UNSURE
<222> (104)..(113)
<223> Xaa is Unknown

<400> 28

Phe Arg Cys Ile Val His Pro Phe Arg Glu Lys Leu Thr Leu Arg Lys

1 5 10 15

Ala Leu Val Thr Ile Ala Val Ile Trp Ala Leu Ala Leu Leu Ile Met
20 25 30

Cys Pro Ser Ala Val Thr Leu Thr Val Thr Arg Glu Glu His His Phe
35 40 45

Met Val Asp Ala Arg Asn Arg Ser Tyr Pro Leu Tyr Ser Cys Trp Glu
50 55 60

Ala Trp Pro Glu Lys Gly Met Arg Arg Val Tyr Thr Thr Val Leu Phe
65 70 75 80

Ser His Ile Tyr Leu Ala Pro Leu Ala Leu Ile Val Val Met Tyr Ala
85 90 95

Arg Ile Ala Arg Lys Leu Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
100 105 110

Xaa Glu Ala Ala Asp Pro Arg Ala Ser Arg Arg Arg Ala Arg Val Val
115 120 125

His Met Leu Val Met Val Ala Leu Phe Phe Thr
130 135

<210> 29
<211> 318
<212> DNA
<213> H.Sapiens

<400> 29
gcacggggcg tgagtcctca ggcacttctt gaggtccttg ttgagcagga agcagacaat 60

tgggttgacg gcagcctggg cgaagctcat ccaaacagca ctggccaggt agccgtgggg 120

cacagcacag gctttcacaa ccactcgcca gtagcaggcc acgatgtagg gtgaccagag 180

gagcagaaag agccgtgtga tcgcgtagaa catgcggccc agctgctttt caccttgac 240

ctcgtccatg cccagtagcc gccggctggc tgcatgccca ttctgccgga tacccagcag 300

ggttggtggc atgggccc 318

<210> 30
<211> 106
<212> PRT
<213> H.Sapiens

<400> 30

Gly Pro Met Pro Pro Thr Leu Leu Gly Ile Arg Gln Asn Gly His Ala
1 5 10 15

Ala Ser Arg Arg Leu Leu Gly Met Asp Glu Val Lys Gly Glu Lys Gln
20 25 30

Leu Gly Arg Met Phe Tyr Ala Ile Thr Leu Leu Phe Leu Leu Leu Trp

35 40 45

Ser Pro Tyr Ile Val Ala Cys Tyr Trp Arg Val Phe Val Lys Ala Cys
50 55 60

Ala Val Pro His Arg Tyr Leu Ala Thr Ala Val Trp Met Ser Phe Ala
65 70 75 80

Gln Ala Ala Val Asn Pro Ile Val Cys Phe Leu Leu Asn Lys Asp Leu
85 90 95

Lys Lys Cys Leu Arg Thr His Ala Pro Cys
100 105

<210> 31
<211> 354
<212> DNA
<213> H.Sapiens

<400> 31
tattctgtaa tgaagaatgt cattcacact gccattggca catccagtgg cctcacctag 60

cattgtgaaa gcccttcggt tggtgtattg ccacttcatt ttaaaaggat gcacaagtcc 120

ctggtgcctt tccacagcaa tgcaggtcat agtgaggatt tctgtcacaa cagcggtaga 180

ctggacaaat ggcaccatct tgcaaatgaa agcacctgca gtaaggaaat aggataaatc 240

atacatcaaa acaaaaagaa taaaggtttc atctgtgtct ttgtaattat cactatcagt 300

ccattctgag cctctgccaa aaagtttgat aattgtaatt actctgtaga caca 354

<210> 32
<211> 117
<212> PRT
<213> H.Sapiens

<400> 32

Val Tyr Arg Val Ile Thr Ile Ile Lys Leu Phe Gly Arg Gly Ser Glu
1 5 10 15

Trp Thr Asp Ser Asp Asn Tyr Lys Asp Thr Asp Glu Thr Phe Ile Leu
20 25 30

Phe Val Leu Met Tyr Asp Leu Ser Tyr Phe Leu Thr Ala Gly Ala Phe
35 40 45

Ile Cys Lys Met Val Pro Phe Val Gln Ser Thr Ala Val Val Thr Glu
50 55 60

Ile Leu Thr Met Thr Cys Ile Ala Val Glu Arg His Gln Gly Leu Val
65 70 75 80

His Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg Arg Ala Phe Thr
85 90 95

Met Leu Gly Glu Ala Thr Gly Cys Ala Asn Gly Ser Val Asn Asp Ile

100 105 110

Leu His Tyr Arg Ile
115

<210> 33
<211> 621
<212> DNA
<213> H.Sapiens

<400> 33
gagcaacatg atctttttga agtacttgac ggtgtcgttc ttgacggtca cgaagcacag 60

agtgttgatc atgctgttgc tcatggcgat gcactcgacg atgtagaagg cagtgaggta 120

gtgcttctcc ttcacaaaca cggtggggaa gaagtcgcgc acgatggtga agccgtagaa 180

gggcgcccag catagcacgt aggcggtgag gatgcacatg agcaccaggc ccgtcttcct 240

gcggcagcgc agcctcttgc ggatctgctc tgtctggaat ccagggaccg ccttgaacca 300

gagctccagg gagatcctgg catagcacag ggtcatggtg accacggggc ccacgaattc 360

tatgccaaag ataaagagga agtaggactt gtagtagagc tgctggtcca caggccagat 420

ctggccgcag aagatctttt cctggctctt gaccatgacg aggacggtct cggtggtgaa 480

gtaggcggaa gggatggcga tcaggatgga cacggtccac accaaggcaa tcaggccagt 540

ggctgtttgg cacttcattc gtggtctcag cggatggaca atagccagat acctagggca 600

agaacacaag tggaggcagc c 621

<210> 34
<211> 207
<212> PRT
<213> H.Sapiens

<400> 34

Gly Cys Leu His Leu Cys Ser Cys Pro Arg Tyr Leu Ala Ile Val His
1 5 10 15

Pro Leu Arg Pro Arg Met Lys Cys Gln Thr Ala Thr Gly Leu Ile Ala
20 25 30

Leu Val Trp Thr Val Ser Ile Leu Ile Ala Ile Pro Ser Ala Tyr Phe
35 40 45

Thr Thr Glu Thr Val Leu Val Ile Val Lys Ser Gln Glu Lys Ile Phe
50 55 60

Cys Gly Gln Ile Trp Pro Val Asp Gln Gln Leu Tyr Tyr Lys Ser Tyr
65 70 75 80

Phe Leu Phe Ile Phe Gly Ile Glu Phe Val Gly Pro Val Val Thr Met
85 90 95

Thr Leu Cys Tyr Ala Arg Ile Ser Arg Glu Leu Trp Phe Lys Ala Val
100 105 110

Pro Gly Phe Gln Thr Glu Gln Ile Arg Lys Arg Leu Arg Cys Arg Arg
115 120 125

Lys Thr Val Leu Val Leu Met Cys Ile Leu Thr Ala Tyr Val Leu Cys
130 135 140

Trp Ala Pro Phe Tyr Gly Phe Thr Ile Val Arg Asp Phe Phe Pro Thr
145 150 155 160

Val Phe Val Lys Glu Lys His Tyr Leu Thr Ala Phe Tyr Ile Val Glu
165 170 175

Cys Ile Ala Met Ser Asn Ser Met Ile Asn Thr Leu Cys Phe Val Thr
180 185 190

Val Lys Asn Asp Thr Val Lys Tyr Phe Lys Lys Ile Met Leu Leu
195 200 205

<210> 35
<211> 483
<212> DNA
<213> H.Sapiens

<400> 35
cagccacact gcagtgatga aatcaaatgt ccaacaccaa ccatagtcac cattactaac 60

taagaagcca caaaacttcc cttccagggt gttcagcagc agggacaggg cccagggcag 120

ggcacacatg acagttgaca ggtttcttgg gcagcagcag cagtaccaga taggccgcag 180

gacagacagg cagcactcag tactgatggc actcagcatg ctcaggccta caaggtaggc 240

aaaggtcatc acgctggtga agaagctagg gaaattgatg gagatggaac agaagaagtt 300

actgaggtac accaggcaat ttataatctg gaagcagagg aagagcaagt cggccccggc 360

caggctgagg acctagacag agaaggcgtt cctgcgcatg cggaagccca ggagccagag 420

cacaaacccg tttcctacca gcccgaccag ggccatgaaa agaatcaggc agacccggat 480

cag 483

<210> 36
<211> 161
<212> PRT
<213> H.Sapiens

<400> 36

Leu Ile Pro Val Phe Leu Ile Leu Phe Ile Ala Leu Val Gly Leu Val
1 5 10 15

Gly Asn Gly Phe Val Leu Trp Leu Leu Gly Phe Arg Met Arg Arg Asn
20 25 30

Ala Phe Ser Val Tyr Val Leu Ser Leu Ala Gly Ala Asp Phe Leu Phe
35 40 45

Leu Cys Phe Gln Ile Ile Asn Cys Leu Val Tyr Leu Ser Asn Phe Phe
50 55 60

Cys Ser Ile Ser Ile Asn Phe Pro Ser Phe Phe Thr Ser Val Met Thr
65 70 75 80

Phe Ala Tyr Leu Val Gly Leu Ser Met Leu Ser Ala Ile Ser Thr Glu
85 90 95

Cys Cys Leu Ser Val Leu Arg Pro Ile Trp Tyr Cys Cys Cys Cys Pro
100 105 110

Arg Asn Leu Ser Thr Val Met Cys Ala Leu Pro Trp Ala Leu Ser Leu
115 120 125

Leu Leu Asn Thr Leu Glu Gly Lys Phe Cys Gly Phe Leu Val Ser Asn
130 135 140

Gly Asp Tyr Gly Trp Cys Trp Thr Phe Asp Phe Ile Thr Ala Val Trp
145 150 155 160

Leu

<210> 37
<211> 330
<212> DNA
<213> H.Sapiens

<400> 37
gagagtctga ttctgactta catcacatat gtaggcctgg gcatttctat ttgcagcctg 60

atcctttgct tgtccgttga ggtcctagtc tggagccaag tgacaaagac agagatcacc 120

tatttacgcc atgtgtgcat tgttaacatt gcagccactt tgctgatggc agatgtgtgg 180

ttcattgtgg ctttcctttct tagtggcccc atcacacacc acaagggatg tgtggcagcc 240

acattttttg gtcatttctt ttacctttct gtatttttct ggatgcttgc caaggcactc 300

cttatcctct atggaatcat gattgttttc 330

<210> 38
<211> 110
<212> PRT
<213> H.Sapiens

<400> 38

Glu Ser Leu Ile Leu Thr Tyr Ile Thr Tyr Val Gly Leu Gly Ile Ser
1 5 10 15

Ile Cys Ser Leu Ile Leu Cys Leu Ser Val Glu Val Leu Val Trp Ser
20 25 30

Gln Val Thr Lys Thr Glu Ile Thr Tyr Leu Arg His Val Cys Ile Val
35 40 45

Asn Ile Ala Ala Thr Leu Leu Met Ala Asp Val Trp Phe Ile Val Ala
50 55 60

Ser Phe Leu Ser Gly Pro Ile Thr His His Lys Gly Cys Val Ala Ala
65 70 75 80

Thr Phe Phe Gly His Phe Phe Tyr Leu Ser Val Phe Phe Trp Met Leu
85 90 95

Ala Lys Ala Leu Leu Ile Leu Tyr Gly Ile Met Ile Val Phe
100 105 110

<210> 39
<211> 628
<212> DNA
<213> H.Sapiens

<400> 39
ttgtgtggca gtagagagat gtcaggcttc agagtcaaca agaactggat ttcaaactgg 60

atttgaggac ccccaccttt ggtaagtgac ttattatctg cgagcctctg tttctctctt 120

ctttaaatga ggacagtaaa tcccatccgg cagggtggtg gggagaatca gagatgatac 180

agctggtgat cacatctggt ttgtgttccc agggcaccca gactagggtt tctgagcatg 240

gatccaaccg tcccagtctt cggtacaaaa ctgacaccaa tcaacggacg tgaggagact 300

ccttgctaca atccgaccct gagcttcacg gtgctgacgt gcatcatttc ccttgtcgga 360

ctgacaggaa accgggtact gctctggctc ctgggctacc gcatgcgcag gaacgctgtc 420

tccatctaca tcctcaacct ggccgcagca gacttcctct tcctcagctt ccagattata 480

cgttcgccat tacgcctcat caatatcagc catctcatcc gcaaaatcct cgtttctgtg 540

atgacctttc cctactttac aggcctgagt atgctgagcg ccatcagcac cgagcgctgc 600

ctgtctgttc tgtggcccat ctggtacc 628

<210> 40
<211> 205
<212> PRT
<213> H.Sapiens

<400> 40

Leu Cys Gly Ser Arg Glu Met Ser Gly Phe Arg Val Asn Lys Asn Trp
1 5 10 15

Ile Ser Asn Trp Ile Gly Pro Pro Pro Leu Val Ser Asp Leu Leu Ser
20 25 30

Ala Ser Leu Cys Phe Ser Leu Leu Met Arg Thr Val Asn Pro Ile Arg
35 40 45

Gln Gly Gly Gly Glu Asn Gln Arg Tyr Ser Trp Ser His Leu Val Cys
    50   55   60

Val Pro Arg Gly Thr Arg Leu Gly Phe Leu Ser Met Asp Pro Thr Val
65   70   75   80

Pro Val Phe Gly Thr Lys Leu Thr Pro Ile Asn Gly Arg Glu Glu Thr
                85   90   95

Pro Cys Tyr Asn Gln Thr Leu Ser Phe Thr Val Leu Thr Cys Ile Ile
            100   105   110

Ser Leu Val Gly Leu Thr Gly Asn Ala Val Val Leu Trp Leu Leu Gly
        115   120   125

Tyr Arg Met Arg Arg Asn Ala Val Ser Ile Tyr Ile Leu Asn Leu Ala
    130   135   140

Ala Ala Asp Phe Leu Phe Leu Ser Phe Gln Ile Ile Arg Ser Pro Leu
145   150   155   160

Arg Leu Ile Asn Ile Ser His Leu Ile Arg Lys Ile Leu Val Ser Val
                165   170   175

Met Thr Phe Pro Tyr Phe Thr Gly Leu Ser Met Leu Ser Ala Ile Ser
            180   185   190

Thr Glu Arg Cys Leu Ser Val Leu Trp Pro Ile Trp Tyr
        195   200   205

<210> 41
<211> 319
<212> DNA
<213> H.Sapiens

<400> 41
acagaaagca aggccaccag gaccttaggc atagtcatgg gagtgtttgt gttgtgctgg 60

ctgcccttct ttgtcttgac gatcacagat cctttcatta attttacaac ccttgaagat 120

ctgtacaatg tcttcctctg gctaggctat ttcaactctg ctttcaatcc cattttatat 180

ggcatgcttt atccttggtt tcgcaaggca ttgaggatga ttgtcacagg catgatcttc 240

caccctgact cttccaccct aagcctcttt tctgccaatg cttaggctgt gttcatcatt 300

caataggact cttttctgg 319

<210> 42
<211> 103
<212> PRT
<213> H.Sapiens

<400> 42

Thr Glu Ser Lys Ala Thr Arg Thr Leu Gly Ile Val Met Gly Val Phe
1   5   10   15

Val Leu Cys Trp Leu Pro Phe Phe Val Leu Thr Ile Thr Asp Pro Phe
20 25 30

Ile Asn Phe Thr Thr Leu Glu Asp Leu Tyr Asn Val Phe Leu Trp Leu
35 40 45

Gly Tyr Phe Asn Ser Ala Phe Asn Pro Ile Leu Tyr Gly Met Leu Tyr
50 55 60

Pro Trp Phe Arg Lys Ala Leu Arg Met Ile Val Thr Gly Met Ile Phe
65 70 75 80

His Pro Asp Ser Ser Thr Leu Ser Leu Phe Ser Ala His Ala Ala Val
85 90 95

Phe Ile Ile Gln Asp Ser Phe
100

<210> 43
<211> 515
<212> DNA
<213> H.Sapiens

<400> 43
taggaatctc agagaagaaa gtaaggaacc agaaaaccat aaaagaatgt aaatggaaaa 60

gaatcagcaa atcttattca cttatcacta aatctaaaat atgtcaaaat acatgaagac 120

aacaaatgct ttagaacaac tgttgaatgt attgtcctac aacttggcat atgatcatgc 180

ttgcctctct atgtccaagt gtttattttt gcagttgacc ttaatttcaa gttagttttg 240

aggtctctac agtaatgttt ttaatctgtc tctacttctt cagaaataa attagttgtt 300

gacaaatcag tccttaagac cttgccgctt acaataagtt ttattgcctt cccaaaccat 360

tggtaaaaga aagcataaat caaggggttc atagctgaat tataataaac acaccaaact 420

aaaatctcat aaacataagg aggagttata aaattcatat aagcatcaat cactgcatca 480

acgaggtatg ctagccaaga gacaagaaat gctgc 515

<210> 44
<211> 148
<212> PRT
<213> H.Sapiens

<400> 44

Leu His Gln Arg Gly Met Val Ala Lys Arg Gln Glu Met Leu Ala Ala
1 5 10 15

Phe Leu Val Ser Trp Leu Pro Tyr Leu Val Asp Ala Val Ile Asp Ala
20 25 30

Tyr Met Asn Phe Ile Thr Pro Pro Tyr Val Tyr Glu Ile Leu Val Trp
35 40 45

Cys Val Tyr Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr Ala Phe Phe
    50   55   60

Tyr Gln Trp Phe Gly Lys Ala Ile Lys Leu Ile Val Ser Gly Lys Val
65   70   75   80

Leu Arg Thr Asp Ser Ser Thr Thr Asn Leu Phe Ser Glu Glu Val Glu
                85   90   95

Thr Asp Lys His Tyr Cys Arg Asp Leu Lys Thr Asn Leu Lys Leu Arg
            100   105   110

Ser Thr Ala Lys Ile Asn Thr Trp Thr Arg Gly Lys His Asp His Met
        115   120   125

Pro Ser Cys Arg Thr Ile His Ser Thr Val Val Leu Lys His Leu Leu
    130   135   140

Ser Ser Cys Ile
145

<210> 45
<211> 726
<212> DNA
<213> H.Sapiens

<400> 45
ctggaaagag gtcctcgatc tatcctctac gccgtccttg gttttggggc tgtgctggca 60

gcgtttggaa acttactggt catgattgct atccttcact tctaacaact gcacacacct 120

acaaactttc tgattgcgtc gctgccctgt gctgacttct tggtgggagt cactgtgatg 180

cccttcagca cagtgaggtc tgtgcagagc tgttggtact ttgggcacag ttactgtaaa 240

ttccatacat cttttgaaac atctttatgt tttgcttctt tatttcattt atgcttgatc 300

tctcttgata gatacattgc tgttactgat cctctgacct atccaaccaa gtttactgtg 360

tcagtttcag ggatatgcat tgttctttcc tggttctttt ctgtcacata cagcttttcg 420

atcttttaca cgggagccac ccaagaagca attgaggaat tagtagttgc tctaacctgt 480

gtaggaggct gccaggctcc actgaatcaa aactgggtcc tactttgttt tcttctattc 540

tttatacaca atgtcgccat ggtgtttata tacagtaaga tatttttggt ggccaagcat 600

caggctagga agatagaaag tacagccagc caagctcagt ccttctcaga gagttacaag 660

gaaagagtag caaaaagaga gagaacgcct gcccaaacct tgggaattgc tatggccgcc 720

tttctt 726

<210> 46
<211> 241
<212> PRT
<213> H.Sapiens

```
<400> 46

Leu Glu Arg Gly Pro Arg Ser Ile Leu Tyr Ala Val Leu Gly Phe Gly
1               5                   10                  15

Ala Val Leu Ala Ala Phe Gly Asn Leu Leu Val Met Ile Ala Ile Leu
            20                  25                  30

His Phe Gln Leu His Thr Pro Thr Asn Phe Leu Ile Ala Ser Leu Ala
        35                  40                  45

Cys Ala Asp Phe Leu Val Gly Val Thr Val Met Pro Phe Ser Thr Val
    50                  55                  60

Arg Ser Val Glu Ser Cys Trp Tyr Phe Gly Asp Ser Tyr Cys Lys Phe
65                  70                  75                  80

His Thr Cys Phe Asp Thr Ser Phe Cys Phe Ala Ser Leu Phe His Leu
                85                  90                  95

Cys Cys Ile Ser Val Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Thr
            100                 105                 110

Tyr Pro Thr Lys Phe Thr Val Ser Val Ser Gly Ile Cys Ile Val Leu
        115                 120                 125

Ser Trp Phe Phe Ser Val Thr Tyr Ser Phe Ser Ile Phe Tyr Thr Gly
    130                 135                 140

Ala Asn Glu Glu Gly Ile Glu Glu Leu Val Val Ala Leu Thr Cys Val
145                 150                 155                 160

Gly Gly Cys Gln Ala Pro Leu Asn Gln Asn Trp Val Leu Leu Cys Phe
                165                 170                 175

Leu Leu Phe Phe Ile Pro Asn Val Ala Met Val Phe Ile Tyr Ser Lys
            180                 185                 190

Ile Phe Leu Val Ala Lys His Gln Ala Arg Lys Ile Glu Ser Thr Ala
        195                 200                 205

Ser Gln Ala Gln Ser Phe Ser Glu Ser Tyr Lys Glu Arg Val Ala Lys
    210                 215                 220

Arg Glu Arg Lys Ala Ala Lys Thr Leu Gly Ile Ala Met Ala Ala Phe
225                 230                 235                 240

Leu


<210>  47
<211>  660
<212>  DNA
<213>  H.Sapiens

<400>  47
aacccggtgg ccttactcct aagaccoctg gccttgtcta tggcctttat caacagctgt      60
```

ctcaatccag ttctctatgt cttcattggg catgacttct gggagcactt gctccactcc 120

ctgctacgag ccttagaacg ggcacttagc gaggagccag atagtgccta aatcccagct 180

cccaggcaga tgagtccttt ataacatgac ccaatttcct actcccattt cccacccactc 240

aatcctcttc ccaaacagct ctaccataat ccaacatcca accgaattta agagaataaa 300

ccacaacttt taagtgagct ctatgtgcta ggtcatgttt tagaatacaa ccttaagtgc 360

ctggaagatg gaggcaagaa acaaacaagg tctcattctt tagaggaaga cagttcacca 420

agactcaaac agaaaaaaag atagttatct tgtgacaaaa caagtcataa aattgggtca 480

ggacctgcag caatgacttt atgctagaat ccagagcact agcaggaaac tgcttaaatt 540

ttacttaatc aaagtcaagt ttggacatac atgtcaggta aaacctagca gagatgagct 600

accttgattt taaaacttca agggatagct caatgtcatc aagatccttt tgatgacttg 660

<210> 48
<211> 211
<212> PRT
<213> H Sapiens

<400> 48

Asn Gln Val Ala Leu Leu Leu Arg Pro Leu Ala Leu Ser Met Ala Phe
1 5 10 15

Ile Asn Ser Cys Leu Asn Pro Val Leu Tyr Val Phe Ile Gly His Asp
20 25 30

Phe Trp Glu His Leu Leu His Ser Leu Leu Ala Ala Leu Glu Arg Ala
35 40 45

Leu Ser Glu Glu Pro Asp Ser Ala Ile Pro Ala Pro Arg Gln Met Ser
50 55 60

Pro Leu His Asp Pro Ile Ser Tyr Ser Ile Phe Pro Pro Leu Asn Pro
65 70 75 80

Leu Pro Lys Gln Leu Tyr His Asn Pro Thr Ser Asn Arg Ile Glu Asn
85 90 95

Lys Pro Gln Leu Leu Ser Glu Leu Tyr Val Leu Gly His Val Leu Glu
100 105 110

Tyr Asn Leu Lys Cys Leu Glu Asp Gly Gly Lys Lys Gln Thr Arg Ser
115 120 125

His Ser Leu Glu Glu Asp Ser Ser Pro Arg Leu Lys Gln Lys Lys Arg
130 135 140

Leu Ser Cys Asp Lys Thr Ser His Lys Ile Gly Ser Gly Pro Ala Ala
145 150 155 160

Met Thr Leu Cys Asn Pro Glu His Gln Glu Thr Ala Ile Leu Leu Asn

```
                    165                 170                 175

Gln Ser Gln Val Trp Thr Tyr Met Ser Gly Lys Thr Gln Arg Ala Thr
            180                 185                 190

Leu Ile Leu Lys Leu Gln Gly Ile Ala Gln Cys His Gln Asp Pro Phe
        195                 200                 205

Asp Asp Leu
    210
```

<210> 49
<211> 465
<212> DNA
<213> H.Sapiens

<400> 49

```
gcttgttcac ggccaccatc ctcaagctgt tgcgcacgga ggaggcgcac ggccgggagc      60

agcggaggcg ggcggtgggc ctggccgcgg tggtcttgct ggcctttgtc acctgcttcg     120

cccccaacaa cttcgtgctc ctggcgcaca tcgtgagccg cctgttctac ggcaagagct     180

actaccacgt gtacaagctc acgctgtgtc tcagctgcct caacaactgt ctggacccgt     240

ttgtttatta ctttgcgtcc cgggaattcc agctgcgcct gcgggaatat ttgggctgcc     300

gccgggtgcc caagagcacc ctggacacgc gccgcgagag cctcttctcc gccaggacca     360

cgtccgtgcg ctccgaggcc ggtgcgcacc ctgaagggat ggagggagcc accaggcccg     420

gcctccagag gcaggagagt gtgttctgag tcccgggggc gcagc                     465
```

<210> 50
<211> 160
<212> PRT
<213> H.Sapiens

<400> 50

```
Leu Phe Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu Glu Ala His
1               5                   10                  15

Gly Arg Glu Gln Arg Arg Arg Ala Val Gly Leu Ala Ala Val Val Leu
            20                  25                  30

Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val Leu Leu Ala
        35                  40                  45

His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr His Val Tyr
    50                  55                  60

Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu Asp Pro Phe
65                  70                  75                  80

Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu Arg Glu Tyr
                85                  90                  95
```

Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr Arg Arg Glu
100 105 110

Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu Ala Gly Ala
115 120 125

His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu Gln Arg Gln
130 135 140

Glu Ser Val Phe Val Pro Gly Ala Gln Ala Ala Pro Pro Gly Leu Arg
145 150 155 160

<210> 51
<211> 603
<212> DNA
<213> H.Sapiens

<400> 51
ttacttattc tgccctttat ccaactttta attccctttg ctattctcct gcctcatttt 60

ctgccctcat tttccctatt atcctgcctc acattgatca agggatgagg ctggcaggat 120

ccggaaccca cagggccccg tgggccatga gaggctcctg gacttgaacc tccggacact 180

cccactctgg ctgccggcag ggatggaagc tggatgagca ggcaggagct ggcagtgggg 240

gtggagagcc ataggctatt ggggtggaca ggcttgggtg cctcatggga gctccccatg 300

ggagctgtgg ccccttgggg cctcttattt ctcaccccag gctttcccgg gagaggttca 360

agtcagaaga tgccccaaag atccacgtgg ccctgggtgg cagcctgttc ctcctgaatc 420

tcgccttctt ggtcaatgtg gggagtgggt cacaccggtc tgatcctgcc tgctgggccc 480

gggggctggt cttccactac ttcctgctct gtgccttcac ctggatgggc cttgaagcct 540

tccacctcta cctgctcgct gtcagggtct tcaacaccta cttcgggcac tacttcctga 600

agc 603

<210> 52
<211> 198
<212> PRT
<213> H.Sapiens

<400> 52

Glu Thr Tyr Ser Ala Leu Tyr Pro Thr Phe Asn Ser Leu Cys Tyr Ser
1 5 10 15

Pro Ala Ser Phe Ser Gly Leu Ile Phe Pro Ile Ile Leu Pro His Ile
20 25 30

Asp Gln Gly Met Arg Leu Ala Gly Ser Gly Thr His Arg Ala Pro Trp
35 40 45

Ala Met Arg Gly Ser Trp Thr Thr Ser Gly His Ser His Ser Gly Cys
50 55 60

Arg Gln Gly Trp Lys Leu Asp Glu Gln Ala Gly Ala Gly Ser Gly Gly
65 70 75 80

Gly Glu Pro Ala Ile Gly Val Asp Arg Leu Gly Cys Leu Met Gly Ala
85 90 95

Pro His Gly Ser Cys Gly Pro Leu Gly Pro Leu Ile Ser His Pro Arg
100 105 110

Leu Ser Arg Glu Arg Phe Lys Ser Glu Asp Ala Pro Lys Ile His Val
115 120 125

Ala Leu Gly Gly Ser Leu Phe Leu Leu Asn Leu Ala Phe Leu Val Asn
130 135 140

Val Gly Ser Gly Ser Lys Gly Ser Asp Ala Ala Cys Trp Ala Arg Gly
145 150 155 160

Ala Val Phe His Tyr Phe Leu Leu Cys Ala Phe Thr Trp Met Gly Leu
165 170 175

Glu Ala Phe His Leu Tyr Leu Leu Ala Val Arg Val Phe Asn Thr Tyr
180 185 190

Phe Gly His Tyr Phe Leu
195

<210> 53
<211> 335
<212> DNA
<213> H.Sapiens

<400> 53
aattggtcgg agagtgcagc tgcttgaaat ggaggattga aatcatcacc aggaggtttc 60

caaacacagc cagcacagcc ccaaagccaa acactatgta cagaatcacc cgggatcccg 120

gcgagaaggg gattttcaca caggacccat tcacgttcgc gtagcacagc tgcacagcca 180

ccagcaggga tgaattgctg ctcataacgc tggtatttac atatggagaa attttgtcct 240

tgttgattat cacaaaaaat acaggattgt tcctgatttt cattgctcct gcggaaaaaa 300

acacatattc accaggatgc cagaggaaat gatca 335

<210> 54
<211> 111
<212> PRT
<213> H.Sapiens

<400> 54

Asp His Phe Leu Trp His Pro Gly Glu Tyr Val Phe Phe Ser Ala Gly
1 5 10 15

Ala Met Lys Ile Arg Asn Asn Pro Val Phe Phe Val Ile Ile Asn Lys
20 25 30

Asp Lys Ile Ser Pro Tyr Val Asn Thr Ser Val Met Ser Ser Asn Ser
35 40 45

Ser Leu Leu Val Ala Val Gln Leu Cys Tyr Ala Asn Val Asn Gly Ser
50 55 60

Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg Val Ile Leu Tyr Ile
65 70 75 80

Val Phe Gly Phe Gly Ala Val Leu Ala Val Phe Gly Asn Leu Leu Val
85 90 95

Met Ile Ser Ile Leu His Phe Lys Gln Leu His Ser Pro Thr Asn
100 105 110

<210> 55
<211> 586
<212> DNA
<213> H.Sapiens

<400> 55
cacatcttaa caagactgaa aaacattgat ttgtttttaa tttgaagagc aatttatttg 60

ctattcattc atagtcttac ttgattttta aaaactcatt tcgcttggta attttaaagc 120

tatcctgaac ttcgtctatc caactgctta tatatgttca gaaaccaaat tcatggttgc 180

tgaactgttc tttaaaacct gaccagttac aataactttt attgctttcc taaaccatgc 240

gtaaaataaa gcataaatca aaggattcat ggctgagtta taataagcac accaacagca 300

tcataaatac aggcaggggt tataaagccc ataaaggcat caattaatga atcaatgcta 360

tatggtaacc atcaaatcat aaatgctacc actgtgaccc ccagggtttt agctgctttt 420

ctctctctcc tgcccactct ggctttgtaa ctctctgagg atgattctgt cttgctacca 480

gtattttcta tctttttcgc ctgtcgtcta gccacaagaa atatgttacc atacagaatt 540

atcataataa aggtaggtat aaagaaggat agaaaatctg tcaaca 586

<210> 56
<211> 190
<212> PRT
<213> H.Sapiens

<400> 56

Leu Thr Asp Phe Leu Ser Phe Phe Ile Pro Thr Phe Ile Met Ile Ile
1 5 10 15

Leu Tyr Gly Asn Ile Phe Leu Val Ala Arg Arg Gln Ala Lys Lys Ile
20 25 30

Glu Asn Thr Gly Ser Lys Thr Glu Ser Ser Ser Glu Ser Tyr Lys Ala
35 40 45

Arg Val Ala Arg Arg Glu Arg Lys Ala Ala Lys Thr Leu Gly Val Thr
    50   55   60

Val Val Ala Phe Met Ile Ser Trp Leu Pro Tyr Ser Ile Asp Ser Leu
65   70   75   80

Ile Asp Ala Phe Met Gly Phe Ile Thr Pro Ala Cys Ile Tyr Glu Ile
                85   90   95

Cys Cys Trp Cys Ala Tyr Tyr Asn Ser Ala Met Asn Pro Leu Ile Tyr
            100   105   110

Ala Leu Phe Tyr Pro Trp Phe Arg Lys Ala Ile Lys Val Ile Val Thr
        115   120   125

Gly Gln Val Leu Lys Asn Ser Ser Ala Thr Met Asn Leu Phe Ser Glu
    130   135   140

His Ile Ala Val Gly Thr Lys Phe Arg Ile Pro Leu Lys Leu Pro Ser
145   150   155   160

Glu Met Ser Phe Lys Ser Ser Lys Thr Met Asn Glu Gln Ile Asn Cys
                165   170   175

Ser Ser Asn Lys Gln Ile Asn Val Phe Gln Ser Cys Asp Val
            180   185   190

<210> 57
<211> 976
<212> DNA
<213> H.Sapiens

<400> 57
tttgtggcaa ggagaccctg atcccggtct tcctgatcct tttcattgcc ctggtcgggc 60

tggtcggaaa cgggtttgtg ctctggctcc tgggcttccg catgcgcatg aacgccttct 120

ctgtctacgt cctcagcctg gccggggccg acttcctctt cctctgcttc cagcttataa 180

attgcctggt gtacctcagt aacttcttct gttccatctc catcaatttc cctagcttcc 240

tcaccactgt gatgacctgt gcctaccttg caggcctgag catgctgagc accgtcagca 300

ccgagcgctg cctgtccgtc ctgtggccca tctggtatcg ctgccgccgc cccagacacc 360

tgtcagcggt cgtgtgtgtc ctgctctggg ccctgtccct actgctgagc atcttggaag 420

ggaagttctg tggcttctta tttagtgatg gtgactctgg ttggtgtcag acatttgatt 480

tcatcactgc agcgtggctg atttttttat tcatggttct ctgtgggtcc agtctggccc 540

tgctggtcag gatcctctgt ggctccaggg gtctgccact gaccaggctg tacctgacca 600

tcctgctcac agtgctggtg tccctcctct gcggcctgcc ctttggcatt cagtggttcc 660

taatattatg gatctggaag gattctgatg tcttattttg tcatattcat ccagtttcag 720

ttgtcctgtc atctcttcac agcagtgccc accccatcat ttacttcttc gtgggctctt 780

ttaggaagca gtgccgggtg cagcacccga tcctcaagct ggctctccag agggctctgc 840

aggacattgc tgaggtggat cacagtgaag gatgcttccg tcagggcacc cggagattca 900

aagaagcatt ctggtgtagg gatggacccc tctacttcca tcatatatat gtggctttga 960

gaggcaactt tgcccc 976

<210> 58
<211> 324
<212> PRT
<213> H.Sapiens

<220>
<221> UNSURE
<222> (266)..(266)
<223> Xaa is Unknown

<400> 58

Cys Gly Lys Glu Thr Leu Ile Pro Val Phe Leu Ile Leu Phe Ile Ala
1 5 10 15

Leu Val Gly Leu Val Gly Asn Gly Phe Val Leu Trp Leu Leu Gly Phe
20 25 30

Arg Met Arg Arg Asn Ala Phe Ser Val Tyr Val Leu Ser Leu Ala Gly
35 40 45

Ala Asp Phe Leu Phe Leu Cys Phe Gln Ile Ile Asn Cys Leu Val Tyr
50 55 60

Leu Ser Asn Phe Phe Cys Ser Ile Ser Ile Asn Phe Pro Ser Phe Phe
65 70 75 80

Thr Thr Val Met Thr Cys Ala Tyr Leu Ala Gly Leu Ser Met Leu Ser
85 90 95

Thr Val Ser Thr Glu Arg Cys Leu Ser Val Leu Trp Pro Ile Trp Tyr
100 105 110

Arg Cys Arg Arg Pro Arg His Leu Ser Ala Val Val Cys Val Leu Leu
115 120 125

Trp Ala Leu Ser Leu Leu Leu Ser Ile Leu Glu Gly Lys Phe Cys Gly
130 135 140

Phe Leu Phe Ser Asp Gly Asp Ser Gly Trp Cys Gln Thr Phe Asp Phe
145 150 155 160

Ile Thr Ala Ala Trp Leu Ile Phe Leu Phe Met Val Leu Cys Gly Ser
165 170 175

Ser Leu Ala Leu Leu Val Arg Ile Leu Cys Gly Ser Arg Gly Leu Pro
180 185 190

Leu Thr Arg Leu Tyr Leu Thr Ile Leu Leu Thr Val Leu Val Ser Leu

```
        195                 200                 205

Leu Cys Gly Leu Pro Phe Gly Ile Gln Trp Phe Leu Ile Leu Trp Ile
    210                 215                 220

Trp Lys Asp Ser Asp Val Leu Phe Cys His Ile His Pro Val Ser Val
225                 230                 235                 240

Val Leu Ser Ser Leu Asn Ser Ser Ala Asn Pro Ile Ile Tyr Phe Phe
                245                 250                 255

Val Gly Ser Phe Arg Lys Gln Trp Arg Xaa Gln His Pro Ile Leu Lys
            260                 265                 270

Leu Ala Leu Gln Arg Ala Leu Gln Asp Ile Ala Glu Val Asp His Ser
        275                 280                 285

Glu Gly Cys Phe Arg Gln Gly Thr Arg Arg Phe Lys Glu Ala Phe Trp
    290                 295                 300

Cys Arg Asp Gly Pro Leu Tyr Phe His His Ile Tyr Val Ala Leu Arg
305                 310                 315                 320

Gly Asn Phe Ala


<210>  59
<211>  578
<212>  DNA
<213>  H.Sapiens

<400>  59
ctttgcatct cactgttgag cagacagcct gctgaaagtt gtcgctgacc accacatata      60

gtaacaggtt accaaaggtg ttcagagcag cataatggtc tacaaacgat gtaagcttca     120

tggatctgat tctcaatgga acaactgatt gaaagcaggc tgagattcga tcctgaatga     180

ccctcaagct atggaagggt aaaaaaccta cgtaaactgc aacaagtagc agaatggtta     240

gcctgctgcc ttttctgcta aggcagctgt cagtttgcag tccatgggtc aaagtgtgga     300

taatcgtggt atagcaaagt gtcactatca ccaaggggag gcagaaagta cttgcagtca     360

aaatcaggtt gtaccactta atagtattga gttcatccca actggtgagg tcgacacagg     420

ctgatctgtt ggtcctgttg gttgatctga tcaagcaagt catcggaatg acaagctaca     480

gtgaaatgat ccacaccacc gcacaggcta caactgcaca tcgagttttg tgaatggaaa     540

agcagctcat tgggtgaatg atcacacagt agcggaag                             578


<210>  60
<211>  192
<212>  PRT
<213>  H.Sapiens

<400>  60
```

Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile His
1   5   10   15

Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile Ser
            20   25   30

Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Thr Asn Arg
        35   40   45

Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu Asn
    50   55   60

Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Ser Thr Phe Cys Leu
65   70   75   80

Pro Leu Val Ile Val Thr Leu Cys Tyr Thr Thr Ile Ile His Thr Leu
                85   90   95

Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ala Arg Arg
            100   105   110

Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro Phe
        115   120   125

His Ile Leu Arg Val Ile Gln Asp Arg Ile Ser Ala Cys Phe Gln Ser
    130   135   140

Val Val Pro Leu Arg Ile Arg Ser Met Lys Leu Thr Ser Phe Leu Asp
145   150   155   160

His Tyr Ala Ala Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val Val
                165   170   175

Val Ser Asp Asn Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys Lys
            180   185   190

<210> 61
<211> 872
<212> DNA
<213> H.Sapiens

<400> 61
gggagggctc gtagacacac taaccctacc ctttctgttt cttcctcatc tttcctttcc 60

atctgtttct catggtctcc tgtctgtctc tctctctctc ccctctttct ctctccctgc 120

tctttctcat ccccttcatt tctgtgtcaa tctcaatcca tttatatcgg tggccacttt 180

tctatctctt tgttctatct ctctctctct ctctttccca ctttgtctct gcacgcctgt 240

tgtgtttttc tgcctgtctc tctcttgccc tcatctctct ctctctctct tgccctcatc 300

tctctgtctc tctgtgtctg tgtctccccc gctcattccc atttgccggt gcaatgtcgc 360

aggacaactc atggagcccc cccgggccca tcgagtaccg gactggctga ccccctaggg 420

ttggcagtag cccctgaccc tcagtatggc caacactacc ggagagcctg aggaggtgag 480

cggcgctctg tccccacogt ccgcatcagc ttatgtgaag ctggtactgc tgggactgat 540

tatgtgcgtg agcctggcgg gtaacgccat cttgtccctg ctggtgctca aggagcgggc 600

cctgcacaag gctccttact acttcctgct ggacctgtgc ctggccgatg gcatacgctc 720

tgccgtctgc ttccccttg tgctggcttc tgtgcgccac ggctcttcat ggaccttcag 720

tgcactcagc tgcaagattg tggcctttat ggccgtgctc ttttgcttcc atgcggcctt 780

catgctgttc tgcatcagcg tcacccgcta catcgccatc gcccaccacc gcttctacgc 840

caagcgcatg acactctgga catgcgcggc tg 872

<210> 62
<211> 143
<212> PRT
<213> H.Sapiens

<400> 62

Met Ala Asn Thr Thr Gly Glu Pro Glu Glu Val Ser Gly Ala Leu Ser
1 5 10 15

Pro Pro Ser Ala Ser Ala Tyr Val Lys Leu Val Leu Leu Gly Leu Ile
20 25 30

Met Cys Val Ser Leu Ala Gly Asn Ala Ile Leu Ser Leu Leu Val Leu
35 40 45

Lys Glu Arg Ala Leu His Lys Ala Pro Tyr Tyr Phe Leu Leu Asp Leu
50 55 60

Cys Leu Ala Asp Gly Ile Arg Ser Ala Val Cys Phe Pro Phe Val Leu
65 70 75 80

Ala Ser Val Arg His Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys
85 90 95

Lys Ile Val Ala Phe Met Ala Val Leu Phe Cys Phe His Ala Ala Phe
100 105 110

Met Leu Phe Cys Ile Ser Val Thr Arg Tyr Met Ala Ile Ala His His
115 120 125

Arg Phe Tyr Ala Lys Arg Met Thr Leu Trp Thr Cys Ala Ala Glu
130 135 140

<210> 63
<211> 962
<212> DNA
<213> H.Sapiens

<400> 63
aaaaattgct gtactgaact attgaatgga acttggaaat aaagtccctt ccaaaataac 60

tattcttcaa cagagagtaa taggtaaatg ttttagaagt gagaggactc aaattgccaa 120

tgatttactc ttttattttt cctcctaggt ttctgggata agtatgtgca aataaaaaat 180

aaacatgaga aggaactgta acctgattat ggatttggga aaaagataag tcaacacaca 240

aagggaaaag taaactgatt gacagccctc aggaatgatg cccttttgcc acaatataat 300

taatatttcc tgtgtgaaaa acaactggtc aaatgatgtc cgtgcttccc tgtacagttt 360

aatggtgctc ataattctga ccacactcgt tggcaatctg atagttattg tttctatatc 420

acacttcaaa caacttcata ccccaacaaa ttggctcatt cattccatgg ccactgtgga 480

ctttcttctg ggtgtctgg tcatgcctta cagtatggtg agatctgctg agcactgttg 540

gtattttgga gaagtcttct gtaaaattca cacaagcacc gacattatgc tgagctcagc 600

ctccattttc catttgtctt tcatctccat tgaccgctac tatgctgtgt gtgatccact 660

gagatataaa gccaagatga atatcttggt tatttgtgtg atgatcttca ttagttggag 720

tgtccctgct gtttttgcat ttggaatgat ctttctggag ctaaacttca aaggcgctga 780

agagatatat tacaaacatg ttcactgcag aggaggttgc tctgtcttct ttagcaaaat 840

atctggggta ctgaccttta tgactttttt tatatacctt ggatctatta tgttatgtgt 900

ctattacaga atatatctta tcgctaaaga acaggcaaga ttaattagtg atgccaatca 960

ga 962

<210> 64
<211> 238
<212> PRT
<213> H.Sapiens

<400> 64

Arg Glu Lys Thr Asp Gln Pro Ser Gly Met Met Pro Phe Cys His Asn
1 5 10 15

Ile Ile Asn Ile Ser Cys Val Lys Asn Asn Trp Ser Asn Asp Val Arg
20 25 30

Ala Ser Leu Tyr Ser Leu Met Val Leu Ile Ile Leu Thr Thr Leu Val
35 40 45

Gly Asn Leu Ile Val Ile Val Ser Ile Ser His Phe Lys Gln Leu His
50 55 60

Thr Pro Thr Asn Trp Leu Ile His Ser Met Ala Thr Val Asp Phe Leu
65 70 75 80

Leu Gly Cys Leu Val Met Pro Tyr Ser Met Val Arg Ser Ala Glu His
85 90 95

Cys Trp Tyr Phe Gly Glu Val Phe Cys Lys Ile His Thr Ser Thr Asp
100 105 110

Ile Met Leu Ser Ser Ala Ser Ile Phe His Leu Ser Phe Ile Ser Ile
115 120 125

Asp Arg Tyr Tyr Ala Val Cys Asp Pro Leu Arg Tyr Lys Ala Lys Met
130 135 140

Asn Ile Leu Val Ile Cys Val Met Ile Phe Ile Ser Trp Ser Val Pro
145 150 155 160

Ala Val Phe Ala Phe Gly Met Ile Phe Leu Glu Leu Asn Phe Lys Gly
165 170 175

Ala Glu Glu Ile Tyr Tyr Lys His Val His Cys Arg Gly Gly Cys Ser
180 185 190

Val Phe Phe Ser Lys Ile Ser Gly Val Leu Thr Phe Met Thr Ser Phe
195 200 205

Tyr Ile Pro Gly Ser Ile Met Leu Cys Val Tyr Tyr Arg Ile Tyr Leu
210 215 220

Ile Ala Lys Glu Gln Ala Arg Leu Ile Ser Asp Ala Asn Gln
225 230 235

<210> 65
<211> 1018
<212> DNA
<213> H.Sapiens

<400> 65
aacagtcccg ggtggaacct gggcatgtat attttgattg ttttatgcat actcctagtg 60

aagaaccaat gtcttgctca gatagaagca agatactcag acttagtttc tctgtagctc 120

ctgcttttta ttattcctgg ttggattgca ccactactca gtttctattt tataatactg 180

attataaaac atggcaggga aataactttg tattggtttt tatggataat ttattatgtg 240

tcctagactc tggccttgtc aaaagaagga cgtcagaagg cacgatgtat tatacttggg 300

aatgatagaa gagactgacc tggtatttcc accggaaga gggaaacgat tttaactaca 360

aatacaggaa tccagcagat ggcatcagag aacactataa aaaagaaacg atttgcaaca 420

gccacctctc ttccaaaaca attccttact tctgtggtct gcaaggcggt tttttgaatg 480

gaacagaaca tagtaatata gcaaaacaca atgatgagaa aagccagcaa gttccacact 540

gttggggaaa agcacacttt taccatctca ggcgtaaaag tcaacagtaa aattactatg 600

gtccaggttg agtatccctt acccaaaatg tttgaaacca gaaatgtttt ggatttcgga 660

tttcggaata tttacacatt cataatgata tatcttggaa atggttccca agtctaaaca 720

caaaatttat ttatgtttca tatacacctt atacacatag tctgaaagta atttgtaca 780

atattttaaa taatttttggg catgaaacaa agtttgcata cattgaacca tcagacagca 840

aaagcttcag gtgtggaatt ttccacttgt ggcatcatgt tgatgctcaa aaagttccat 900

attttagagc atttcaaatt ttggattttc aaattacaaa tgcttaacct gtacttagat 960

gttaaataca gtgcctcttc cacgggcact ttcaggaagc attcttttat ataagccc 1018

<210> 66
<211> 327
<212> PRT
<213> H.Sapiens

<400> 66

```
Tyr Ile Lys Glu Cys Phe Leu Lys Val Pro Val Glu Glu Ala Leu Tyr
1               5                   10                  15

Leu Thr Ser Lys Tyr Arg Leu Ser Ile Cys Asn Leu Lys Ile Gln Asn
            20                  25                  30

Leu Lys Cys Ser Lys Ile Trp Asn Phe Leu Ser Ile Asn Met Met Pro
        35                  40                  45

Gln Val Glu Asn Ser Thr Pro Glu Ala Phe Ala Val Trp Phe Asn Val
    50                  55                  60

Cys Lys Leu Cys Phe Met Pro Lys Ile Ile Asn Ile Val Gln Asn Tyr
65                  70                  75                  80

Phe Gln Thr Met Cys Ile Arg Cys Ile Asn Ile Asn Lys Phe Cys Val
                85                  90                  95

Thr Trp Glu Pro Phe Pro Arg Tyr Ile Ile Met Asn Val Ile Phe Arg
            100                 105                 110

Asn Pro Lys Ser Lys Thr Phe Leu Val Ser Asn Ile Leu Gly Lys Gly
        115                 120                 125

Tyr Ser Thr Cys Thr Thr Val Ile Leu Leu Leu Thr Phe Thr Pro Glu
    130                 135                 140

Met Leu Lys Val Cys Phe Ser Pro Thr Gly Val Asn Leu Leu Ala Phe
145                 150                 155                 160

Leu Ile Ile Val Phe Ser Tyr Ile Thr Met Phe Cys Ser Ile Gln Lys
                165                 170                 175

Thr Ala Leu Gln Thr Thr Glu Val Arg Asn Cys Phe Gly Arg Glu Val
            180                 185                 190

Ala Val Ala Asn Arg Phe Phe Phe Ile Val Phe Ser Asp Ala Ile Cys
        195                 200                 205

Trp Ile Pro Val Phe Val Val Lys Ile Leu Ser Leu Phe Arg Val Glu
    210                 215                 220

Ile Pro Gly Gln Ser Leu Leu Ser Phe Pro Ser Ile Ile His Arg Ala
225                 230                 235                 240

Phe Leu Arg Pro Ser Phe Asp Lys Ala Arg Val Asp Thr Ile Ile His
```

245 250 255

Lys Asn Gln Tyr Lys Val Ile Ser Leu Pro Cys Phe Ile Ile Ser Ile
260 265 270

Ile Lys Lys Leu Ser Ser Gly Ala Ile Gln Pro Gly Ile Ile Lys Ser
275 280 285

Arg Ser Tyr Arg Glu Thr Lys Ser Glu Tyr Leu Ala Ser Ile Ala Arg
290 295 300

His Trp Phe Phe Thr Arg Ser Met His Lys Thr Ile Lys Ile Tyr Met
305 310 315 320

Pro Arg Phe His Pro Gly Leu
325

<210> 67
<211> 1251
<212> DNA
<213> H.Sapiens

<400> 67
actaccatgg aagctgaccc ggtgccacct ggccacagcc cccgcacaga gcttgatgat 60

gaggactccc acccccaagg tggctgggac acggtcttcc tggtggccct gctgctcctt 120

gggctgccag ccaatgggtt gatggcgtgg ctggccggct cccaggcccg gcatggagct 180

ggcacgcgtc tggcgctgct cctgctcagc ctggccctct ctcacttctt gttcctggca 240

gcagcggcct tccagctccc acagatcccg catcggggac actggccccc ggggaccgct 300

gcctgccgct tctacttact cctatggggc ctgtcctact cctccggccc cttcctgctg 360

gcggccctca gcctcgaccg ctgcctgctg gcgctgtgcc cacactggta ccctgggcac 420

cgcccagtcc gcctgcccct ctgggtctgc gccggtgtct gggtgctggc cacactcttc 480

agcgtgccct ggctggtctt ccccgaggct gccgtctggt ggtacgacct ggtcatctgc 540

ctggacttct gggacagcga ggagctgtcg ctgaggatgc tggaggtcct ggggggcttc 600

ctgcctttcc tcctgctgct cgtctgccac gtgctcaccc aggccacagc ctgtcgcacc 660

tgccaccgcc aacagcagcc cgcagcctgc cggggcttcg cccgtgtggc caggaccatt 720

ctgtcagcct atgtggtcct gaggctgccc taccagctgg cccagctgct ctacctggcc 780

ttcctgtggg acgtctactc tggctacctg ctctgggagg ccctggtcta ctccgactac 840

ctgatcctac tcaacagctg cctcagcccc ttcctctgcc tcatggccag tgccgacctc 900

cggaccctgc tgcgctccgt gctctcgtcc ttcgcggcag ctctctgcga ggagcggccg 960

ggcagcttca cgcccactga gccacagacc cagctagatt ctgagggtcc aactctgcca 1020

gagccgatgg cagaggccca gtcacagatg gatcctgtgg cccagcctca ggtgaacccc 1080

acactccagc cacgatggga tcccccagct cagccacagc tgaaccctac ggcccagcca 1140

cagtcggatc ccacagccca gccacagctg aacctcatgg cccagcccca gtcagattct 1200

gtggcccagc cacaggcaga cactaacgtc cagacccctg cacctgctgc c 1251

<210> 68
<211> 417
<212> PRT
<213> H.Sapiens

<400> 68

Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro Arg Thr
1 5 10 15

Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp Thr Val
20 25 30

Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro Ala Asn Gly Leu Met
35 40 45

Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr Arg Leu
50 55 60

Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe Leu Ala
65 70 75 80

Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His Trp Pro
85 90 95

Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly Val Ser
100 105 110

Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp Arg Cys
115 120 125

Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro Val Arg
130 135 140

Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr Leu Phe
145 150 155 160

Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp Tyr Asp
165 170 175

Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser Leu Arg
180 185 190

Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe Leu Leu Leu Leu Val
195 200 205

Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg Thr Cys His Arg Gln
210 215 220

Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg Thr Ile
225 230 235 240

Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr Gln Leu Ala Gln Leu
245 250 255

Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu Leu Trp
260 265 270

Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser Cys Leu
275 280 285

Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr Leu Leu
290 295 300

Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Glu Glu Arg Pro
305 310 315 320

Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser Glu Gly
325 330 335

Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met Asp Pro
340 345 350

Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln Pro Arg Ser Asp Pro
355 360 365

Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser Asp Pro
370 375 380

Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser Asp Ser
385 390 395 400

Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala Pro Ala
405 410 415

Ala

<210> 69
<211> 659
<212> DNA
<213> H.Sapiens

<400> 69
tacaggcctg agcatgctgg gctccatcag caccaagcac tgccctgcca tcctgtggcc 60

catctagtac cgctgccacc accccacaca cctgtcagca gtcgtgtgtc ctgctctggg 120

ccctgtccct gctgcagagc atcctggaat ggatgttctg tggcttcctg tctagtggtg 180

ctgattctgt ttgctgtgaa acctcagatt tcatcacagt cacatggctg atttttttat 240

gtgtggttct ctgcgggtcc agcccggttc tgctggtcag gatcctttgt ggatcccgga 300

agatgccctt gaccaggctg tacatgacca tcctgctcag agtgctggtc ttcctcctct 360

gtgacctgcc ctttggcatt cagtgattcc tatttttctg gatccacgtg gatttgtcac 420

gttcgtctag tttccatttt cctgtccact cttaacagca gtgccaaccc cattatttac 480

ttcttcatgg gctcctttag gcagcttcac aacaggaaga ctctctagct ggttctccag 540

agggctctgc aggacacgcc tgaggtggaa gaaggcagat ggcggctttc tgaggaaacc 600

ctggagctgt catgaagcag attggggcca tgaggaagag cctctgccct gtcagtcag 659

<210> 70
<211> 213
<212> PRT
<213> H.Sapiens

<400> 70

```
Tyr Arg Pro Glu His Ala Gly Leu His Gln His Gln Ala Leu Pro Val
1               5                   10                  15

His Pro Val Ala His Leu Val Pro Leu Pro Pro Pro His Thr Pro Val
            20                  25                  30

Ser Ser Arg Val Ser Cys Ser Gly Pro Cys Pro Cys Cys Arg Ala Ser
        35                  40                  45

Trp Asn Gly Cys Ser Val Ala Ser Cys Leu Val Val Leu Ile Leu Phe
    50                  55                  60

Gly Val Lys His Gln Ile Ser Ser Gln Ser His Gly Phe Phe Tyr Val
65                  70                  75                  80

Trp Phe Ser Ala Gly Pro Ala Arg Phe Cys Trp Ser Gly Ser Phe Val
                85                  90                  95

Asp Pro Gly Arg Cys Pro Pro Gly Cys Thr Pro Ser Cys Ser Glu Cys
            100                 105                 110

Trp Ser Ser Ser Ser Val Thr Cys Pro Leu Ala Phe Ser Asp Ser Tyr
        115                 120                 125

Phe Ser Gly Ser Thr Trp Ile Cys His Val Arg Leu Val Ser Ile Phe
    130                 135                 140

Leu Ser Thr Leu Asn Ser Ser Ala Asn Pro Ile Ile Tyr Phe Phe Met
145                 150                 155                 160

Gly Ser Phe Arg Gln Leu Gln Asn Arg Lys Thr Leu Leu Val Leu Gln
                165                 170                 175

Arg Ala Leu Gln Asp Thr Pro Glu Val Glu Glu Gly Arg Trp Arg Leu
            180                 185                 190

Ser Glu Glu Thr Leu Glu Leu Ser Ser Arg Leu Gly Pro Gly Arg Ala
        195                 200                 205

Ser Ala Leu Ser Val
    210
```

<210> 71
<211> 559
<212> DNA
<213> H.Sapiens

<400> 71
atgccgaagg caggccgcag aagaagaagg gaggacggtg aggaggatga gcccagggaa 60

gcccgggggt ggggcccgct ggggccctcg ctccaccggc agcagcagca tcaggctggc 120

cccacccatg gtgcaacaca gcagagccag cagcaccgct gccaccagcc accgcgtccg 180

gcacaagtgg cggctgggct ccccgaagaa ctgggtgcag gcgccgctga gcagcagctg 240

cagcagcagg cagagggccc aggtgagggc gcacacacag gtggtcaggt ggcgtgggcg 300

gcggccacag taccaggctg ggaagacggc ggccaggcac tgctccacgc tcacggccgc 360

caggacactc agccccacga tgtagcagaa gaagcgcagc gttgccagcc tcgtctgcac 420

gaagccgggg aagtccagcc ggccttgcag caagtcgggg acgatggcca ccatgtggca 480

gccaggaagg atgagatccg cgcaggccac gtccaggagg tagatggcga aagggtttct 540

gtagacattg gagctgagc 559

<210> 72
<211> 213
<212> PRT
<213> H.Sapiens

<400> 72

Leu Ser Ser Asn Val Tyr Arg Asn Pro Phe Ala Ile Tyr Leu Leu Asp
1 5 10 15

Val Ala Cys Ala Asp Leu Ile Phe Leu Gly Cys His Met Val Ala Ile
20 25 30

Val Pro Asp Leu Leu Gln Gly Arg Leu Asp Phe Pro Gly Phe Val Gln
35 40 45

Thr Ser Leu Ala Thr Leu Arg Phe Phe Cys Tyr Ile Val Gly Leu Ser
50 55 60

Leu Leu Ala Ala Val Ser Val Glu Gln Cys Leu Ala Ala Leu Phe Pro
65 70 75 80

Ala Trp Tyr Ser Cys Arg Arg Pro Arg His Leu Thr Thr Cys Val Cys
85 90 95

Ala Leu Thr Trp Ala Leu Cys Leu Leu Leu His Leu Thr Thr Cys Val
100 105 110

Cys Ala Leu Thr Trp Ala Leu Cys Leu Leu Leu His Leu Leu Leu Ser
115 120 125

Gly Ala Cys Thr Leu Leu Leu Ser Gly Ala Cys Thr Gln Phe Phe Gly
130 135 140

Glu Pro Ser Arg His Leu Cys Arg Thr Leu Trp Leu Val Ala Ala Val
145 150 155 160

Leu Leu Ala Leu Leu Cys Cys Thr Met Cys Gly Ala Ser Leu Met Leu
                165   170   175

Leu Leu Arg Val Glu Arg Gly Pro Gln Arg Pro Pro Pro Arg Gly Phe
            180   185   190

Pro Gly Leu Ile Leu Leu Thr Val Leu Leu Phe Ser Ser Ala Ala Cys
        195   200   205

Leu Arg His
    210

<210> 73
<211> 1008
<212> DNA
<213> H.Sapiens

<400> 73
atggaatcat ctttctcatt tggagtgatc cttgctgtcc tggcctccct catcattgct 60

actaacacac tagtggctgt ggctgtcctg ctgttgatcc acaagaatga tggtgtcagt 120

ctctgcttca ccttgaatct ggctgtggct gacaccttga ttggtgtggc catctctggc 180

ctactcacag accagctctc cagcccttct cggcccacac agaagaccct gtgcagcctg 240

cggatggcat ttgtcacttc ctccgcagct gcctctgtcc tcacggtcat gctgatcacc 300

tttgacaggt accttgccat caagcagccc ttccgctact tgaagatcat gagtgggttc 360

gtggccgggg cctgcattgc cgggctgtgg ttagtgtctt acctcattgg cttcctccca 420

ctcggaatcc ccatgttcca gcagactgcc tacaaagggc agtgcagctt ctttgctgta 480

tttcaccctc acttcgtgct gaccctctcc tgcgttggct tcttcccagc catgctcctc 540

tttgtcttct tctactgcga catgctcaag attgcctcca tgcacagcca gcagattcga 600

aagatggaac atgcaggagc catggctgga ggttatcgat ccccacggac tcccagcgac 660

ttcaaagctc tccgtactgt gtctgttctc attgggagct ttgctctatc ctggaccccc 720

ttccttatca ctggcattgt gcaggtggcc tgccaggagt gtcacctcta cctagtgctg 780

gaacggtacc tgtggctgct cggcgtgggc aactccctgc tcaacccact catctatgcc 840

tattggcaga aggaggtgcg actgcagctc taccacatgg ccctaggagt gaagaaggtg 900

ctcacctcat tcctcctctt tctctcggcc aggaattgtg gcccagagag gcccagggaa 960

agttcctgtc acatcgtcac tatctccagc tcagagtttg atggctaa 1008

<210> 74
<211> 335
<212> PRT
<213> H.Sapiens

<400> 74

```
Met Glu Ser Ser Phe Ser Phe Gly Val Ile Leu Ala Val Leu Ala Ser
1               5                   10                  15

Leu Ile Ile Ala Thr Asn Thr Leu Val Ala Val Ala Val Leu Leu Leu
            20                  25                  30

Ile His Lys Asn Asp Gly Val Ser Leu Cys Phe Thr Leu Asn Leu Ala
        35                  40                  45

Val Ala Asp Thr Leu Ile Gly Val Ala Ile Ser Gly Leu Leu Thr Asp
    50                  55                  60

Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu
65                  70                  75                  80

Arg Met Ala Phe Val Thr Ser Ser Ala Ala Ala Ser Val Leu Thr Val
                85                  90                  95

Met Leu Ile Thr Phe Asp Arg Tyr Leu Ala Ile Lys Gln Pro Phe Arg
            100                 105                 110

Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly
        115                 120                 125

Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro
    130                 135                 140

Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val
145                 150                 155                 160

Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro
                165                 170                 175

Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala
            180                 185                 190

Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met
        195                 200                 205

Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu
    210                 215                 220

Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro
225                 230                 235                 240

Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu
                245                 250                 255

Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser
            260                 265                 270

Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu
        275                 280                 285

Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe
    290                 295                 300
```

Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu
305 310 315 320

Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly
325 330 335

<210> 75
<211> 2137
<212> DNA
<213> H.Sapiens

<400> 75
aactggaagg gcagccgtct gccgcccacg aacaccttct caagcacttt gagtgaccac 60

ggcttgcaag ctggtggctg gccccccgag tcccgggctc tgaggcacgg ccgtcgactt 120

aagcgttgca tcctgttacc tggagaccct ctgagctctc acctgctact tctgccgctg 180

cttctgcaca gagcccgggc gaggacccct ccaggatgca ggtcccgaac agcacgggcc 240

cggacaacgc gacgctgcac atgctgcgga acccggcgat cgcggtggcc ctgcccgtgg 300

tgtactcgct ggtggcggcg gtcagcatcc cgggcaacct cttctctctg tgggtgctgt 360

gccggcgcat ggggcccaga tccccgtcgg tcatcttcat gatcaacctg agcgtcacgg 420

acctgatgct ggccagcgtg ttgcctttcc aaatctacta ccattgcaac cgccaccact 480

gggtattcgg cgtgctgctt tgcaacgtcg tgacgtggcc cttttacgca aacatgtatt 540

ccagcatcct caccatgacc tgtatcagcg tggagcgctt cctgggggtc ctgtacccgc 600

tcagctccaa gcgctggcgc cgccgtcgtt acgcggtggc cgcgtgtgca gggacctggc 660

tgctgctcct gaccgccctg tccccgctgg cgcgcaccga tctcacctac ccggtgcacg 720

ccctgggcat catcacctgc ttcgacgtcc tcaagtggac gatgctcccc agcgtggcca 780

tgtgggccgt gttcctcttc accatcttca tcctgctgtt cctcatcccg ttcgtgatca 840

ccgtggcttg ttcacggccc accatcctca agctgttgcg cacggaggag gcgcacggcc 900

gggagcagcg gaggcgcgcg gtgggcctgg ccgcggtggt cttgctggcc tttgtcacct 960

gcttcgcccc caacaacttc gtgctcctgg cgcacatcgt gagccgcctg ttctacggca 1020

agagctacta ccacgtgtac aagctcacgc tgtgtctcag ctgcctcaac aactgtctgg 1080

acccgtttgt ttattacttt gcgtcccggg aattccagct gcgcctgcgg gaatatttgg 1140

gctgccgccg ggtgcccaga gacaccctgg acacgcgccg cgagagcctc ttctccgcca 1200

ggaccacgtc cgtgcgctcc gaggccggtg cgcaccctga agggatggag ggagccacca 1260

ggcccggcct ccagaggcag gagagtgtgt tctgagtccc gggggcgcag cttggagagc 1320

cgggggcgca gcttggagga tccaggggcg catggagagg ccacggtgcc agaggttcag 1380

ggagaacagc tgcgttgctc ccaggcactg cagaggcccg gtggggaggg gtctccaggc 1440

tttattcctc ccaggcactg cagaggcacc ggtgaggaag ggtctccagg cttcactcag 1500

ggtagagaaa caagcaaagc ccagcagcgc acagggtgct tgttatcctg cagagggtgc 1560

ctctgcctct ctgtgtcagg ggacagcttg tgtcaccccg cccggctaat tttgtatttt 1620

ttttagtag agctgggctg tcacccccga gctccttaga cactcctcac acctgtccat 1680

acccgaggat ggatattcaa ccagccccac cgcctacccg actcggtttc tggatctcct 1740

ctgtgggcga actgcgagcc ccattcccag ctcttctccc tgctgacatc gtcccttagc 1800

acacctgtcc atacccgagg atggatattc aaccagcccc accgcctacc cgactcggtt 1860

tctggatatc ctctgtgggc gaactgcgag ccccattccc agctcttctc cctgctgaca 1920

tcgtccctta gttgtggttc tggccttctc cattctcctc caggggttct ggtctccgta 1980

gcccggtgca cgccgaaatt tctgtttatt tcactcaggg gcactgtggt tgctgtggtt 2040

ggaattcttc tttcagagga gcgcctgggg ctcctgcaag tcagctactc tccgtgccca 2100

cttcccctca cacacacacc cccctcgtgc cgaattc 2137

<210> 76
<211> 359
<212> PRT
<213> H.Sapiens

<400> 76

Met Gln Val Pro Asn Ser Thr Gly Pro Asp Asn Ala Thr Leu Gln Met
1   5   10   15

Leu Arg Asn Pro Ala Ile Ala Val Ala Leu Pro Val Val Tyr Ser Leu
            20   25   30

Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Ser Leu Trp Val Leu
        35   40   45

Cys Arg Arg Met Gly Pro Arg Ser Pro Ser Val Ile Phe Met Ile Asn
    50   55   60

Leu Ser Val Thr Asp Leu Met Leu Ala Ser Val Leu Pro Phe Gln Ile
65   70   75   80

Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Leu Leu Cys
                85   90   95

Asn Val Val Thr Val Ala Phe Tyr Ala Asn Met Tyr Ser Ser Ile Leu
            100   105   110

Thr Met Thr Cys Ile Ser Val Glu Arg Phe Leu Gly Val Leu Tyr Pro
        115   120   125

Leu Ser Ser Lys Arg Trp Arg Arg Arg Arg Tyr Ala Val Ala Ala Cys
    130   135   140

Ala Gly Thr Trp Leu Leu Leu Leu Thr Ala Leu Ser Pro Leu Ala Arg
145 150 155 160

Thr Asp Leu Thr Tyr Pro Val His Ala Leu Gly Ile Ile Thr Cys Phe
165 170 175

Asp Val Leu Lys Trp Thr Met Leu Pro Ser Val Ala Met Trp Ala Val
180 185 190

Phe Leu Phe Thr Ile Phe Ile Leu Leu Phe Leu Ile Pro Phe Val Ile
195 200 205

Thr Val Ala Cys Tyr Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu
210 215 220

Glu Ala His Gly Arg Glu Gln Arg Arg Arg Ala Val Gly Leu Ala Ala
225 230 235 240

Val Val Leu Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val
245 250 255

Leu Leu Ala His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr
260 265 270

His Val Tyr Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu
275 280 285

Asp Pro Phe Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu
290 295 300

Arg Glu Tyr Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr
305 310 315 320

Arg Arg Glu Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu
325 330 335

Ala Gly Ala His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu
340 345 350

Gln Arg Gln Glu Ser Val Phe
355

<210> 77
<211> 1197
<212> DNA
<213> H.Sapiens

<400> 77
atggagtcgg ggctgctgcg gccgccgccg gtgagcgagg tcatcgtcct gcattacaac 60

tacaccggca agctccgcgg tgcgcgctac cagccgggtg ccggcctgcg cgccgacgcc 120

gtggtgtgcc tggcggtgtg cgccttcatc gtgctagaga atctagccgt gttgttggtg 180

ctcggacgcc acccgcgctt ccacgctccc atgttcctgc tcctgggcag cctcacgttg 240

tcggatctgc tggcaggcgc cgcctacgcc gccaacatcc tactgtcggg gccgctcacg 300

ctgaaactgt ccccggcgct ctggttcgca cgggagggag gcgtcttcgt ggcactcact 360
gcgtccgtgc tgagcctcct ggccatcgcg ctggagcgca gcctcaccat ggcgcgcagg 420
ggccccgcgc ccgtctccag tcgggggcgc acgctggcga tggcagccgc ggcctggggc 480
gtgtcgctgc tcctcgggct cctgccagcg ctgggctggc attgcctggg tcgcctggac 540
gcttgctcca ctgtcttgcc gctctacgcc aaggcctacg tgctcttctg cgtgctcgcc 600
ttcgtgggca tcctggccgc tatctgtgca ctctacgcgc gcatctactg ccaggtacgc 660
gccaacgcgc ggcgcctgcc ggcacggccc gggactgcgg ggaccacctc gacccgggcg 720
cgtcgcaagc cgcgctcgct ggccttgctg cgcacgctca gcgtggtgct cctggccttt 780
gtggcatgtt ggggccccct cttcctgctg ctgttgctcg acgtggcgtg ccccgcgcgc 840
acctgtcctg tactcctgca ggccgatccc ttcctgggac tggccatggc caactcactt 900
ctgaacccca tcatctacac gctcaccaac cgcgacctgc gccacgcgct cctgcgcctg 960
gtctgctgcg gacgccactc ctgcggcaga gacccgagtg gctcccagca gtcggcgagc 1020
gcggctgagg cttccggggg cctgcgccgc tgcctgcccc cgggccttga tgggagcttc 1080
agcggctcgg agcgctcctc gccccagcgc gacgggctgg acaccagcgg ctccacaggc 1140
agccccggtg cacccacagc cgcccggact ctggtatcag aacccgctgc agactga 1197

<210> 78
<211> 398
<212> PRT
<213> H.Sapiens

<400> 78

Met Glu Ser Gly Leu Leu Arg Pro Ala Pro Val Ser Glu Val Ile Val
1 5 10 15

Leu His Tyr Asn Tyr Thr Gly Lys Leu Arg Gly Ala Arg Tyr Gln Pro
20 25 30

Gly Ala Gly Leu Arg Ala Asp Ala Val Val Cys Leu Ala Val Cys Ala
35 40 45

Phe Ile Val Leu Glu Asn Leu Ala Val Leu Leu Val Leu Gly Arg His
50 55 60

Pro Arg Phe His Ala Pro Met Phe Leu Leu Leu Gly Ser Leu Thr Leu
65 70 75 80

Ser Asp Leu Leu Ala Gly Ala Ala Tyr Ala Ala Asn Ile Leu Leu Ser
85 90 95

Gly Pro Leu Thr Leu Lys Leu Ser Pro Ala Leu Trp Phe Ala Arg Glu
100 105 110

Gly Gly Val Phe Val Ala Leu Thr Ala Ser Val Leu Ser Leu Leu Ala
115 120 125

Ile Ala Leu Glu Arg Ser Leu Thr Met Ala Arg Arg Gly Pro Ala Pro
130 135 140

Val Ser Ser Arg Gly Arg Thr Leu Ala Met Ala Ala Ala Ala Trp Gly
145 150 155 160

Val Ser Leu Leu Leu Gly Leu Leu Pro Ala Leu Gly Trp Asn Cys Leu
165 170 175

Gly Arg Leu Asp Ala Cys Ser Thr Val Leu Pro Leu Tyr Ala Lys Ala
180 185 190

Tyr Val Leu Phe Cys Val Leu Ala Phe Val Gly Ile Leu Ala Ala Ile
195 200 205

Cys Ala Leu Tyr Ala Arg Ile Tyr Cys Gln Val Arg Ala Asn Ala Arg
210 215 220

Arg Leu Pro Ala Arg Pro Gly Thr Ala Gly Thr Thr Ser Thr Arg Ala
225 230 235 240

Arg Arg Lys Pro Arg Ser Leu Ala Leu Leu Arg Thr Leu Ser Val Val
245 250 255

Leu Leu Ala Phe Val Ala Cys Trp Gly Pro Leu Phe Leu Leu Leu Leu
260 265 270

Leu Asp Val Ala Cys Pro Ala Arg Thr Cys Pro Val Leu Leu Gln Ala
275 280 285

Asp Pro Phe Leu Gly Leu Ala Met Ala Asn Ser Leu Leu Asn Pro Ile
290 295 300

Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg His Ala Leu Leu Arg Leu
305 310 315 320

Val Cys Cys Gly Arg His Ser Cys Gly Arg Asp Pro Ser Gly Ser Gln
325 330 335

Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly Gly Leu Arg Arg Cys Leu
340 345 350

Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly Ser Glu Arg Ser Ser Pro
355 360 365

Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser Thr Gly Ser Pro Gly Ala
370 375 380

Pro Thr Ala Ala Arg Thr Leu Val Ser Glu Pro Ala Ala Asp
385 390 395

<210> 79
<211> 1041
<212> DNA
<213> H.Sapiens

<400> 79
atgtacaacg ggtcgtgctg ccgcatcgag ggggacacca tctcccaggt gatgccgccg 60

ctgctcattg tggcctttgt gctgggcgca ctaggcaatg gggtcgccct gtgtggtttc 120

tgcttccaca tgaagacctg gaagcccagc actgtttacc ttttcaattt ggccgtggct 180

gatttcctcc ttatgatctg cctgcctttt cggacagact attacctcag acgtagacac 240

tgggcttttg gggacattcc ctgccgagtg gggctcttca cgttggccat gaacagggcc 300

gggagcatcg tgttccttac ggtggtggct gcggacaggt atttcaaagt ggtccacccc 360

caccacgcgg tgaacactat ctccacccgg gtggcggctg gcatcgtctg caccctgtgg 420

gccctggtca tcctgggaac agtgtatctt ttgctggaga accatctctg cgtgcaagag 480

acggccgtct cctgtgagag cttcatcatg gagtcggcca atggctggca tgacatcatg 540

ttccagctgg agttctttat gcccctcggc atcatcttat tttgctcctt caagattgtt 600

tggagcctga ggcggaggca gcagctggcc agacaggctc ggatgaagaa ggcgacccgg 660

ttcatcatgg tggtggcaat tgtgttcatc acatgctacc tgcccagcgt gtctgctaga 720

ctctatttcc tctggacggt gccctcgagt gcctgcgatc cctctgtcca tggggccctg 780

cacataaccc tcagcttcac ctacatgaac agcatgctgg atcccctggt gtattatttt 840

tcaagcccct cctttcccca attctacaac aagctcaaaa tctgcagtct gaaacccaag 900

cagccaggac actcaaaaac acaaaggccg gaagagatgc caatttcgaa cctcggtcgc 960

aggagttgca tcagtgtggc aaatagtttc caaagccagt ctgatggcca atgggatccc 1020

cacattgttg agtggcactg a 1041

<210> 80
<211> 346
<212> PRT
<213> H.Sapiens

<400> 80

Met Tyr Asn Gly Ser Cys Cys Arg Ile Glu Gly Asp Thr Ile Ser Gln
1 5 10 15

Val Met Pro Pro Leu Leu Ile Val Ala Phe Val Leu Gly Ala Leu Gly
20 25 30

Asn Gly Val Ala Leu Cys Gly Phe Cys Phe His Met Lys Thr Trp Lys
35 40 45

Pro Ser Thr Val Tyr Leu Phe Asn Leu Ala Val Ala Asp Phe Leu Leu
50 55 60

Met Ile Cys Leu Pro Phe Arg Thr Asp Tyr Tyr Leu Arg Arg Arg His
65 70 75 80

Trp Ala Phe Gly Asp Ile Pro Cys Arg Val Gly Leu Phe Thr Leu Ala
85 90 95

Met Asn Arg Ala Gly Ser Ile Val Phe Leu Thr Val Val Ala Ala Asp
100 105 110

Arg Tyr Phe Lys Val Val His Pro His His Ala Val Asn Thr Ile Ser
115 120 125

Thr Arg Val Ala Ala Gly Ile Val Cys Thr Leu Trp Ala Leu Val Ile
130 135 140

Leu Gly Thr Val Tyr Leu Leu Leu Glu Asn His Leu Cys Val Gln Glu
145 150 155 160

Thr Ala Val Ser Cys Glu Ser Phe Ile Met Glu Ser Ala Asn Gly Trp
165 170 175

His Asp Ile Met Phe Gln Leu Glu Phe Phe Met Pro Leu Gly Ile Ile
180 185 190

Leu Phe Cys Ser Phe Lys Ile Val Trp Ser Leu Arg Arg Arg Gln Gln
195 200 205

Leu Ala Arg Gln Ala Arg Met Lys Lys Ala Thr Arg Phe Ile Met Val
210 215 220

Val Ala Ile Val Phe Ile Thr Cys Tyr Leu Pro Ser Val Ser Ala Arg
225 230 235 240

Leu Tyr Phe Leu Trp Thr Val Pro Ser Ser Ala Cys Asp Pro Ser Val
245 250 255

His Gly Ala Leu His Ile Thr Leu Ser Phe Thr Tyr Met Asn Ser Met
260 265 270

Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ser Pro Ser Phe Pro Lys Phe
275 280 285

Tyr Asn Lys Leu Lys Ile Cys Ser Leu Lys Pro Lys Gln Pro Gly His
290 295 300

Ser Lys Thr Gln Arg Pro Glu Glu Met Pro Ile Ser Asn Leu Gly Arg
305 310 315 320

Arg Ser Cys Ile Ser Val Ala Asn Ser Phe Gln Ser Gln Ser Asp Gly
325 330 335

Gln Trp Asp Pro His Ile Val Glu Trp His
340 345

<210> 81
<211> 2525
<212> DNA
<213> H.Sapiens

<400> 81
caagaatgac aggtgacttc ccaagtatgc ctggccacaa tacctccagg aattcctctt 60

ggcatcctat agtgacaccc cacttaatca gcctctactt catagtgctt attggcgggc 120

tggtgggtgt catttccatt cttttcctcc tggtgaaaat gaacacccgg tcagtgacca 180

ccatggcggt cattaacttg gtggtggtcc acagcgtttt tctgctgaca gtgccctttc 240

gcttgaccta cctcatcaag aagacttgga tgtttgggct gcccttctgc aaattgtgca 300

gtgccatgct gcacatccac atgtacctca cgttcctatt ctatgtggtg atcctggtca 360

ccagatacct catcttcttc aagtgtaaag acaaagtgga attctacaga aaactgcatg 420

ctgtggctgc cagtgctggc atgtggacgc tggtgattgt cattgtggta cccctggttg 480

tctcccggta tggaatccat gaggaataca atgaggagca ctgttttaaa tttcacaaag 540

agcttgctta cacatatgtg aaaatcatca actatatgat agtcattttt gtcatagccg 600

ttgctgtgat tctgttggtc ttccaggtct tcatcattat gttgatggtg cagaagctac 660

gccactcttt actatcccac caggagttct gggctcagct gaaaaaccta ttttttatag 720

ggtcatccct tgtttgtttc cttccctacc acttctttag gatctattac ttgaatgttg 780

tgacgcattc caatgcctgt aacagcaagg ttgcatttta taacggaatc ttcttgagtg 840

taacagcaat tagctgctat gatttgcttc tctttgtctt tgggggaagc cattggttta 900

agcaaaagat aattggctta tggaattgtg ttttgtgccg ttagccacaa actacagtat 960

tcatatttgc ttcctttata ttgggactac aaatgggtat aggggaggta agaatggtat 1020

ttcattactt gatcaaaacc atgccttgat gtacccaaaa caaaaggact ataaaatgca 1080

agagccctca ttgtagtcct tatgggatcc ctcccatctc tgagtgatgg ccgtccaaag 1140

accagtgttg ttgaatccac ctggagttgc aatattacat tattttccag tacagaatgt 1200

ctgtgtggcc catgaaagca acctaggttt taagagtttt agagtttcat tagctcattc 1260

taagttcctc tgtttgaagc atggtctctt aggtttggga ctgaactgag acctttagtt 1320

cttttcatcc cacttcacct taggtaagta aattctggcc accacccagc tctaaagacc 1380

caaactctcc ttcgctaacc aggttagatg tcccattcat ctcatgccct gataaaaact 1440

gataagggga gagaatagtt aaaaattttt ctagggtatc ataactctgg taggaagtca 1500

tctgtctaga aatcaagaga aaaagaacgt gtggcctcct gttataacaa gggtttctag 1560

atttgtcctg tgaaaggtcg tttaaggact tggggatcaa cttcctcaat tatcacccat 1620

tccactgttg ctccaaaaat catttaaaag cttactggac atatctacat aatggtgaaa 1680

ctgtaattta gagactatcc ctgactaatg tgctggtagg cattaaaatg agttcccaag 1740

ggaagtgatt aaaatttttt tctcttctgt tttttgagag aatttctaga tgtcctgggc 1800

cacagttaat taagattttt aggggggaca gaaagttata ctgaaatctt tagagctccc 1860

ttccgccgtt aaaattatat atatatatat ttaaattata ccttaagttc tggggtacat 1920

gtgcagaatg tgcaggtttg ttacataggt atacacgtgc catggtggtt tgcggcacct 1980

gtcaacccat ctacattagg tatttctcct aatgctctcc ctcccctagc cccccacccc 2040

tggacaggcc ccattgtgtg atgttcccct ccctgtgtcc atgtgttttc attgttcaac 2100

tcccacttct aagtgagaac atgcggtgtt tggttttctg ttcctgtgtt agtttgctga 2160

gaatgatggt ttccaggtta aaattatata tttttaaata aatgaaaact gtgtttttaa 2220

aagaggactt ttgagaagta tatagaaaaa ccattaattt agactctgtg agattaggtt 2280

gcctgaagaa gttttctgaa atatttgaag agtggataaa taaatgtccc ccaaagcaat 2340

aaaatcataa tcctttaaaa tataggaaaa ataactaatg ggaactaggc ttaatacttg 2400

ggatgaaata atctgtacaa caaactccca tgacacatgt ttacctatgt aacaaacctg 2460

cacatgtacc cctgaactta aaataaaatt taaagtataa taataaaata atatggattt 2520

tcttt 2525

<210> 82
<211> 312
<212> PRT
<213> H.Sapiens

<400> 82

Met Thr Gly Asp Phe Pro Ser Met Pro Gly His Asn Thr Ser Arg Asn
1 5 10 15

Ser Ser Cys Asp Pro Ile Val Thr Pro His Leu Ile Ser Leu Tyr Phe
20 25 30

Ile Val Leu Ile Gly Gly Leu Val Gly Val Ile Ser Ile Leu Phe Leu
35 40 45

Leu Val Lys Met Asn Thr Arg Ser Val Thr Thr Met Ala Val Ile Asn
50 55 60

Leu Val Val Val His Ser Val Phe Leu Leu Thr Val Pro Phe Arg Leu
65 70 75 80

Thr Tyr Leu Ile Lys Lys Thr Trp Met Phe Gly Leu Pro Phe Cys Lys
85 90 95

Phe Val Ser Ala Met Leu His Ile His Met Tyr Leu Thr Phe Leu Phe
100 105 110

Tyr Val Val Ile Leu Val Thr Arg Tyr Leu Ile Phe Phe Lys Cys Lys
115 120 125

Asp Lys Val Glu Phe Tyr Arg Lys Leu His Ala Val Ala Ala Ser Ala

```
130                 135                 140

Gly Met Trp Thr Leu Val Ile Val Ile Val Val Pro Leu Val Val Ser
145                 150                 155                 160

Arg Tyr Gly Ile His Glu Glu Tyr Asn Glu Glu His Cys Phe Lys Phe
                165                 170                 175

His Lys Glu Leu Ala Tyr Thr Tyr Val Lys Ile Ile Asn Tyr Met Ile
            180                 185                 190

Val Ile Phe Val Ile Ala Val Ala Val Ile Leu Leu Val Phe Gln Val
        195                 200                 205

Phe Ile Ile Met Leu Met Val Gln Lys Leu Arg His Ser Leu Leu Ser
    210                 215                 220

His Gln Glu Phe Trp Ala Gln Leu Lys Asn Leu Phe Phe Ile Gly Val
225                 230                 235                 240

Ile Leu Val Cys Phe Leu Pro Tyr Gln Phe Phe Arg Ile Tyr Tyr Leu
                245                 250                 255

Asn Val Val Thr His Ser Asn Ala Cys Asn Ser Lys Val Ala Phe Tyr
            260                 265                 270

Asn Glu Ile Phe Leu Ser Val Thr Ala Ile Ser Cys Tyr Asp Leu Leu
        275                 280                 285

Leu Phe Val Phe Gly Gly Ser His Trp Phe Lys Gln Lys Ile Ile Gly
    290                 295                 300

Leu Trp Asn Cys Val Leu Cys Arg
305                 310

<210>  83
<211>  1125
<212>  DNA
<213>  H.Sapiens

<400>  83
gcaggagcac tgaaaatcag gaacaatcct gtattttttg tgataatcaa caaggacaaa      60

acttctccat atgtaaataa cagcattatg agcagcaatt catcctgcct ggtggctgtg     120

cagctgtgct acgcgaacgt gaatgggtcc tgtgtgaaaa tccccttctc gccgggatcc     180

cgggtgattc tgtacatagt gtttggcttt ggggctgtgc tggctgtgtt tggaaacctc     240

ctggtgatga tttcaatcct ccatttcaag cagctgcact ctccgaccaa ttttctcgtt     300

gcctctctgg cctgcgctgc tttcttggtg ggtgtgactg tcatgccctt cagcatggtc     360

aggacggtgg aaagctgctg gtattttggg aggagttttt gtactttcca cacctgctgt     420

gatgtggcat tttgttactc ttctctcttt cacttgtgct tcatctccat cgacaggtac     480

attgcggtta ctgaccccct ggtctatcct accaagttca ccgtatctgt gtcaggaatt     540
```

tgcatcagcg tgtcctggat cctgcccctc atgtacagcg gtgctgtctt ctacacaggt 600

gtctatgacg atgggctgga ggaattatct gatgccctaa actgtatagg aggttgtcag 660

acggttgtaa atcaaaactg ggtgttgaca gattttctat cctttcttat acctaccttt 720

attatgataa ttctgtatgg taacatattt cttgtggcta gacgacaggc gaaaaagata 780

gaaaatactg gtagcaagac agaatcatcc tcagagagtt acaaagccag agtggccagg 840

agagagagaa aagcagctaa aaccctgggg gtcacagtgg tagcatttat gatttcatgg 900

ttaccatata gcattgattc attaattgat gcctttatgg gctttataac ccctgcctgt 960

atttatgaga tttgctgttg gtgtgcttat tataactcag ccatgaatcc tttgatttat 1020

gctttatttt acccatggtt taggaaagca ataaaagtta ttgtaactgg tcaggtttta 1080

aagaacagtt cagcaaccat gaatttgttt tctgaacata tataa 1125

<210> 84
<211> 345
<212> PRT
<213> H.Sapiens

<400> 84

Met Ser Ser Asn Ser Ser Leu Leu Val Ala Val Gln Leu Cys Tyr Ala
1 5 10 15

Asn Val Asn Gly Ser Cys Val Lys Ile Pro Phe Ser Pro Gly Ser Arg
20 25 30

Val Ile Leu Tyr Ile Val Phe Gly Phe Gly Ala Val Leu Ala Val Phe
35 40 45

Gly Asn Leu Leu Val Met Ile Ser Ile Leu His Phe Lys Gln Leu His
50 55 60

Ser Pro Thr Asn Phe Leu Val Ala Ser Leu Ala Cys Ala Asp Phe Leu
65 70 75 80

Val Gly Val Thr Val Met Pro Phe Ser Met Val Arg Thr Val Glu Ser
85 90 95

Cys Trp Tyr Phe Gly Arg Ser Phe Cys Thr Phe His Thr Cys Cys Asp
100 105 110

Val Ala Phe Cys Tyr Ser Ser Leu Phe His Leu Cys Phe Ile Ser Ile
115 120 125

Asp Arg Tyr Ile Ala Val Thr Asp Pro Leu Val Tyr Pro Thr Lys Phe
130 135 140

Thr Val Ser Val Ser Gly Ile Cys Ile Ser Val Ser Trp Ile Leu Pro
145 150 155 160

Leu Met Tyr Ser Gly Ala Val Phe Tyr Thr Gly Val Tyr Asp Asp Gly

```
                165                 170                 175

Leu Glu Glu Leu Ser Asp Ala Leu Asn Cys Ile Gly Gly Cys Gln Thr
            180                 185                 190

Val Val Asn Gln Asn Trp Val Leu Thr Asp Phe Leu Ser Phe Phe Ile
        195                 200                 205

Pro Thr Phe Ile Met Ile Ile Leu Tyr Gly Asn Ile Phe Leu Val Ala
    210                 215                 220

Arg Arg Gln Ala Lys Lys Ile Glu Asn Thr Gly Ser Lys Thr Glu Ser
225                 230                 235                 240

Ser Ser Glu Ser Tyr Lys Ala Arg Val Ala Arg Arg Glu Arg Lys Ala
                245                 250                 255

Ala Lys Thr Leu Gly Val Thr Val Val Ala Phe Met Ile Ser Trp Leu
            260                 265                 270

Pro Tyr Ser Ile Asp Ser Leu Ile Asp Ala Phe Met Gly Phe Ile Thr
        275                 280                 285

Pro Ala Cys Ile Tyr Glu Ile Cys Cys Trp Cys Ala Tyr Tyr Asn Ser
    290                 295                 300

Ala Met Asn Pro Leu Ile Tyr Ala Leu Phe Tyr Pro Trp Phe Arg Lys
305                 310                 315                 320

Ala Ile Lys Val Ile Val Thr Gly Gln Val Leu Lys Asn Ser Ser Ala
                325                 330                 335

Thr Met Asn Leu Phe Ser Glu His Ile
            340                 345

<210>  85
<211>  1020
<212>  DNA
<213>  H.Sapiens

<400>  85
accatgaatg agccactgga ctatttagca aatgcttctg atttccccga ttatgcagct      60

gcttttggaa attgcactga tgaaaacatc ccactcaaga tgcactacct ccctgttatt     120

tatggcatta ttttcctcgt gggatttcca ggcaatgcag tagtgatatc cacttacatt     180

ttcaaaatga gaccttggaa gagcagcacc atcattatgc tgaacctggc ctgcacagat     240

ctgctgtatc tgaccagcct ccccttcctg attcactact atgccagtgg cgaaaactgg     300

atctttggag atttcatgtg taagtttatc cgcttcagct tccatttcaa cctgtatagc     360

agcatcctct tcctcacctg tttcagcatc ttccgctact gtgtgatcat tcacccaatg     420

agctgctttt ccattcacaa aactcgatgt gcagttgtag cctgtgctgt ggtgtggatc     480

atttcactgg tagctgtcat tccgatgacc ttcttgatca catcaaccaa caggaccaac     540
```

agatcagcct gtctcgacct caccagttcg gatgaactca atactattaa gtggtacaac 600

ctgattttga ctgccagtac tttctgcctc cccttggtga tagtgacact ttgctatacc 660

acgattatcc acactttgac ccatggactg caaactgaca gctgccttaa gcagaaagca 720

cgaaggctaa ccattctgct actccttgca ttttacgtat gttttttacc cttccatatc 780

ttgagggtca ttcaggatcg aatctcagcc tgctttcaat cagttgttcc attgagaatc 840

agatccatga agcttacatc gtttctagac cattatgctg ctctgaacac ctttggtaac 900

ctgttactat atgtggtggt cagcgacaac tttcagcagg ctgtctgctc aacagtgaga 960

tgcaaagtaa gcgggaacct tgagcaagca aagaaaatta gttactcaaa caacccttga 1020

<210> 86
<211> 336
<212> PRT
<213> H.Sapiens

<400> 86

Met Asn Glu Pro Leu Asp Tyr Leu Ala Asn Ala Ser Asp Phe Pro Asp
1 5 10 15

Tyr Ala Ala Ala Phe Gly Asn Cys Thr Asp Glu Asn Ile Pro Leu Lys
20 25 30

Met His Tyr Leu Pro Val Ile Tyr Gly Ile Ile Phe Leu Val Gly Phe
35 40 45

Pro Gly Asn Ala Val Val Ile Ser Thr Tyr Ile Phe Lys Met Arg Pro
50 55 60

Trp Lys Ser Ser Thr Ile Ile Met Leu Asn Leu Ala Cys Thr Asp Leu
65 70 75 80

Leu Tyr Leu Thr Ser Leu Pro Phe Leu Ile His Tyr Tyr Ala Ser Gly
85 90 95

Glu Asn Trp Ile Phe Gly Asp Phe Met Cys Lys Phe Ile Arg Phe Ser
100 105 110

Phe His Phe Asn Leu Tyr Ser Ser Ile Leu Phe Leu Thr Cys Phe Ser
115 120 125

Ile Phe Arg Tyr Cys Val Ile Ile His Pro Met Ser Cys Phe Ser Ile
130 135 140

His Lys Thr Arg Cys Ala Val Val Ala Cys Ala Val Val Trp Ile Ile
145 150 155 160

Ser Leu Val Ala Val Ile Pro Met Thr Phe Leu Ile Thr Ser Thr Asn
165 170 175

Arg Thr Asn Arg Ser Ala Cys Leu Asp Leu Thr Ser Ser Asp Glu Leu
180 185 190

Asn Thr Ile Lys Trp Tyr Asn Leu Ile Leu Thr Ala Ser Thr Phe Cys
195 200 205

Leu Pro Leu Val Ile Val Thr Leu Cys Tyr Thr Thr Ile Ile His Thr
210 215 220

Leu Thr His Gly Leu Gln Thr Asp Ser Cys Leu Lys Gln Lys Ala Arg
225 230 235 240

Arg Leu Thr Ile Leu Leu Leu Leu Ala Phe Tyr Val Cys Phe Leu Pro
245 250 255

Phe His Ile Leu Arg Val Ile Gln Asp Arg Ile Ser Ala Cys Phe Gln
260 265 270

Ser Val Val Pro Leu Arg Ile Arg Ser Met Lys Leu Thr Ser Phe Leu
275 280 285

Asp His Tyr Ala Ala Leu Asn Thr Phe Gly Asn Leu Leu Leu Tyr Val
290 295 300

Val Val Ser Asp Asn Phe Gln Gln Ala Val Cys Ser Thr Val Arg Cys
305 310 315 320

Lys Val Ser Gly Asn Leu Glu Gln Ala Lys Lys Ile Ser Tyr Ser Asn
325 330 335

<210> 87
<211> 1138
<212> DNA
<213> H.Sapiens

<400> 87
aaaaattgct gtactgaact attgaatgga acttggaaat aaagttccct cccaaataac 60

tattcttcaa cagagagtaa taggtaaatg ttttagaagt gagaggactc aaattgccaa 120

tgatttactc ttttattttt cctcctaggt ttctgggata agtatgtgca aataaaaaat 180

aaacatgaga aggaactgta acctgattat ggatttggga aaaagataaa tcaacacaca 240

aagggaaaag taaactgatt gacagccctc aggaatgatg cccttttgcc acaatataat 300

taatatttcc tgtgtgaaaa acaactggtc aaatgatgtc cgtgcttccc tgtacagttt 360

aatggtgctt ataattctga ccacactggt tggcaatctg atagttattg tttctatatc 420

acacttcaaa caacttcata ccccaacaaa ttggctcatt cattccatgg ccactgtgga 480

ctttcttctg ggtgtctggg tcatgcctta cagtatggtg agatctgctg agcactgttg 540

gtattttgga gaagtcttct gtaaaattca cacaagcacc gacattatgc tgagctcagc 600

ctccattttc catttgtctt tcatctccat tgacaactac tatgctgtgt gtgatccact 660

gagatataaa gccaagatga atatcttggt tatttgtgtg atgatcttca ttagttggag 720

tgtccctgct gtttttgcat ttggaatgat ctttctggag ctaaacttca aaggcgctga 780

```
agagatatat tacaaacatg ttcactgcag aggaggttgc tctgtcttct ttagcaaaat    840
atctggggta ctgaccttta tgacttcttt ttatatacct ggatctatta tgttatgtgt    900
ctattacaga atatatctta tggctaaaga acaggcaaga ttaattagtg atgccaatca    960
gaagctccaa attggattgg aaatgaaaaa tggaatttca caaagcaaag aaaggaaagc   1020
tgtgaagaca ttgggcattg tgatgggagt ttttctaata tgctggtgcc ctttctttat   1080
ctgtacagtc atggaccctt ttcttcacta cattattcca cctactttga atgatgta     1138
```

<210> 88
<211> 296
<212> PRT
<213> H.Sapiens

<400> 88

```
Met Met Pro Phe Cys His Asn Ile Ile Asn Ile Ser Cys Val Lys Asn
1               5                   10                  15

Asn Trp Ser Asn Asp Val Arg Ala Ser Leu Tyr Ser Leu Met Val Leu
            20                  25                  30

Ile Ile Leu Thr Thr Leu Val Gly Asn Leu Ile Val Ile Val Ser Ile
        35                  40                  45

Ser His Phe Lys Gln Leu His Thr Pro Thr Asn Trp Leu Ile His Ser
    50                  55                  60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser
65                  70                  75                  80

Met Val Arg Ser Ala Glu His Cys Trp Tyr Phe Gly Glu Val Phe Cys
                85                  90                  95

Lys Ile His Thr Ser Thr Asp Ile Met Leu Ser Ser Ala Ser Ile Phe
            100                 105                 110

His Leu Ser Phe Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro
        115                 120                 125

Leu Arg Tyr Lys Ala Lys Met Asn Ile Leu Val Ile Cys Val Met Ile
    130                 135                 140

Phe Ile Ser Trp Ser Val Pro Ala Val Phe Ala Phe Gly Met Ile Phe
145                 150                 155                 160

Leu Glu Leu Asn Phe Lys Gly Ala Glu Glu Ile Tyr Tyr Lys His Val
                165                 170                 175

His Cys Arg Gly Gly Cys Ser Val Phe Phe Ser Lys Ile Ser Gly Val
            180                 185                 190

Leu Thr Phe Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Leu Cys
        195                 200                 205
```

Val Tyr Tyr Arg Ile Tyr Leu Ile Ala Lys Glu Gln Ala Arg Leu Ile
    210   215   220

Ser Asp Ala Asn Gln Lys Leu Gln Ile Gly Leu Glu Met Lys Asn Gly
225   230   235   240

Ile Ser Gln Ser Lys Glu Arg Lys Ala Val Lys Thr Leu Gly Ile Val
                245   250   255

Met Gly Val Phe Leu Ile Cys Trp Cys Pro Phe Phe Ile Cys Thr Val
            260   265   270

Met Asp Pro Phe Leu His Tyr Ile Ile Pro Pro Thr Leu Asn Asp Ala
        275   280   285

Arg Gly Ser Arg Ala Asn Ser Ala
    290   295

<210> 89
<211> 1023
<212> DNA
<213> H.Sapiens

<400> 89
ggaatgatgc ccttttgcca caatataatt aatatttcct gtgtgaaaaa caactgatca 60

aatgatgtcc gtgcttccct gtacagttta atggtgctca taattctgac cacactcgtt 120

ggcaatctga tagttattgt ttctatatca cacttcaaac aacttcatac cccaacaaat 180

tggctcattc atttcatggc cactgtggac tttcttctgg ggtgtctggt catgccttac 240

agtatggtga gatctgctga gcactgttgg tattttggag aagtcttctg taaaattcac 300

acaagcaccg acattatgct gagctcagcc tccattttcc atttgtcttt catctccatt 360

gaccgctact atgctgtgtg tgatccactg agatataaag ccaagatgaa tatcttggtt 420

atttgtgtga tgatcttcat tagttggagt gtccctgctg tttttgcatt tggaatgatc 480

tttctggagc tgaacttcaa aggcgctgaa gagatatatt acaaacatgt tcactgcaga 540

ggaggttgct ctgtcttctt tagcaaaata tctggggtac tgacctttat gacttctttt 600

tatatacctg gatctattat gttatgtgtc tattacagaa tatatcttat cgctaaagaa 660

caggcaagat taattagtga tgccaatcag aagctccaaa ttggattgga aatgaaaaat 720

ggaatttcac aaagcaaaga aaggaaagct gtgaagacat tggggattgt gatgggagtt 780

ttcctaatat gctggtgccc tttctttatc tgtacagtca tggacccttt tcttcactac 840

attattccac ctactttgaa tgatgtattg attgggtttg gctacttgaa ctctacattt 900

aatccaatgg tttatgcatt tttctatcct tggtttagaa aagcactgaa gatgatgctg 960

tttggtaaaa ttttccaaaa agattcatcc aggtgtaaat tatttttgga attgagttca 1020

tag 1023

<210> 90
<211> 339
<212> PRT
<213> H.Sapiens

<400> 90

Met Met Pro Phe Cys His Asn Ile Ile Asn Ile Ser Cys Val Lys Asn
1   5   10   15

Asn Trp Ser Asn Asp Val Arg Ala Ser Leu Tyr Ser Leu Met Val Leu
            20   25   30

Ile Ile Leu Thr Thr Leu Val Gly Asn Leu Ile Val Ile Val Ser Ile
        35   40   45

Ser His Phe Lys Gln Leu His Thr Pro Thr Asn Trp Leu Ile His Ser
    50   55   60

Met Ala Thr Val Asp Phe Leu Leu Gly Cys Leu Val Met Pro Tyr Ser
65   70   75   80

Met Val Arg Ser Ala Glu His Cys Trp Tyr Phe Gly Glu Val Phe Cys
                85   90   95

Lys Ile His Thr Ser Thr Asp Ile Met Leu Ser Ser Ala Ser Ile Phe
            100   105   110

His Leu Ser Phe Ile Ser Ile Asp Arg Tyr Tyr Ala Val Cys Asp Pro
        115   120   125

Leu Arg Tyr Lys Ala Lys Met Asn Ile Leu Val Ile Cys Val Met Ile
    130   135   140

Phe Ile Ser Trp Ser Val Pro Ala Val Phe Ala Phe Gly Met Ile Phe
145   150   155   160

Leu Glu Leu Asn Phe Lys Gly Ala Glu Glu Ile Tyr Tyr Lys His Val
                165   170   175

His Cys Arg Gly Gly Cys Ser Val Phe Phe Ser Lys Ile Ser Gly Val
            180   185   190

Leu Thr Phe Met Thr Ser Phe Tyr Ile Pro Gly Ser Ile Met Leu Cys
        195   200   205

Val Tyr Tyr Arg Ile Tyr Leu Ile Ala Lys Glu Gln Ala Arg Leu Ile
    210   215   220

Ser Asp Ala Asn Gln Lys Leu Gln Ile Gly Leu Glu Met Lys Asn Gly
225   230   235   240

Ile Ser Gln Ser Lys Glu Arg Lys Ala Val Lys Thr Leu Gly Ile Val
                245   250   255

Met Gly Val Phe Leu Ile Cys Trp Cys Pro Phe Phe Ile Cys Thr Val

```
                260                 265                 270

Met Asp Pro Phe Leu His Tyr Ile Ile Pro Pro Thr Leu Asn Asp Val
        275                 280                 285

Leu Ile Trp Phe Gly Tyr Leu Asn Ser Thr Phe Asn Pro Met Val Tyr
    290                 295                 300

Ala Phe Phe Tyr Pro Trp Phe Arg Lys Ala Leu Lys Met Met Leu Phe
305                 310                 315                 320

Gly Lys Ile Phe Gln Lys Asp Ser Ser Arg Cys Lys Leu Phe Leu Glu
            325                 330                 335

Leu Ser Ser


<210>  91
<211>  1696
<212>  DNA
<213>  H.Sapiens

<400>  91
ctgtaaagta gattgtatga ggactccatg aggtcatcca cttcaagtcc ttggcatagc      60

ataattactc aaaaggtgat gacaatggcg caggggaggga tggtgacttg cctggagatc     120

cacagcaccg tctctcccat actcggtcat tcacaccatc attgattcac caggcaccac     180

tccgtgtcca gcaggactct ggggacccca aatggacact accatggaag ctgacctggg     240

tgccactggc cacaggcccc gcacagagct tgatgatgag gactcctacc cccaaggtgc     300

ctgggacacg gtcttcctgg tggccctgct gctcctgggg ctgccagcca atgggttgat     360

ggcgtggctg gccggctccc aggccccggca tggagctggc acgcgtctgg cgctgctcct     420

gctcagcctg gccctctctg acttcttgtt cctggcagca gcggccttcc agatcctaga     480

gatccggcat gggggacact ggccgctggg gacagctgcc tgccgcttct actacttcct     540

atggggcgtg tcctactcct ccggcctctt cctgctggcc gccctcagcc tcgaccgctg     600

cctgctggcg ctgtgcccac actggtaccc tgggcaccgc ccagtccgcc tgccctctg      660

ggtctgcgcc ggtgtctggg tgctggccac actcttcagc gtgccctggc tggtcttccc     720

cgaggctgcc gtctggtggt acgacctggt catctgcctg gacttctggg acagcgagga     780

gctgtcgctg aggatgctgg aggtcctggg gggcttcctg cctttcctcc tgctgctcgt     840

ctgccacgtg ctcacccagg ccacagcctg tcgcacctgc cacccgccac agcagcccgc     900

agcctgccgg ggcttcgccc gtgtggccag gaccattctg tcagcctatg tggtcctgag     960

gctgccctac cagctggccc agctgctcta cctggccttc ctgtgggacg tctactctgg    1020

ctacctgctc tgggaggccc tggtctactc cgactacctg atcctactca acagctgcct    1080
```

cagcccctte ctctgcctca tggccagtgc cgacctccgg accctgctgc gctccgtgct 1140

ctcgtccttc ggggcagctc tctgcgagga gcggccgggc agcttcacgc ccactgagcc 1200

acagacccag ctagattctg agggtccaac tctgccagag ccgatggcag aggcccagtc 1260

acagatggat cctgtggccc agcctcaggt gaaccccaca ctccagccac gatcggatcc 1320

cacagctcag ccacagctga accctacggc ccagccacag tcggatccca cagcccagcc 1380

acagctgaac ctcatggccc agccacagtc agattctgtg gcccagccac aggcagacac 1440

taacgtccag accctgcac ctgctgccag ttctgtgccc agtccctgtg atcragcttc 1500

cccaacccca tcctcgcatc ctaccccagg ggcccttgag gacccagcca caccttcctgc 1560

ctctgaagga gaaagcccca gcagcacccc gccagaggcg gccccgggcg caggccccac 1620

gtgggggtcc aggcacacgc aggcccacca gagcagtgaa agagcccagg gcagacagag 1680

gaaccagcca gtcaga 1696

<210> 92
<211> 505
<212> PRT
<213> H Sapiens

<400> 92

Leu Ala Trp Arg Cys Thr Ala Pro Ser Leu Pro Tyr Ser Val Ile His
1 5 10 15

Thr Ile Ile Asp Ser Pro Gly Thr Thr Pro Cys Pro Ala Gly Leu Trp
20 25 30

Gly Pro Gln Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly
35 40 45

His Arg Pro Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly
50 55 60

Gly Trp Asp Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro
65 70 75 80

Ala Asn Gly Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly
85 90 95

Ala Gly Thr Arg Leu Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp
100 105 110

Phe Leu Phe Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His
115 120 125

Gly Gly His Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe
130 135 140

Leu Trp Gly Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu
145 150 155 160

```
Ser Leu Asp Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly
                165                 170                 175

His Arg Pro Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val
            180                 185                 190

Leu Ala Thr Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala
        195                 200                 205

Val Trp Trp Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu
    210                 215                 220

Glu Leu Ser Leu Arg Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe
225                 230                 235                 240

Leu Leu Leu Leu Val Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg
                245                 250                 255

Thr Cys His Arg Gln Gln Glu Pro Ala Ala Cys Arg Gly Phe Ala Arg
            260                 265                 270

Val Ala Arg Thr Ile Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr
        275                 280                 285

Gln Leu Ala Glu Leu Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser
    290                 295                 300

Gly Tyr Leu Leu Trp Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu
305                 310                 315                 320

Leu Asn Ser Cys Leu Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp
                325                 330                 335

Leu Arg Thr Leu Leu Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu
            340                 345                 350

Cys Glu Glu Arg Pro Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Glu
        355                 360                 365

Leu Asp Ser Glu Gly Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln
    370                 375                 380

Ser Gln Met Asp Pro Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln
385                 390                 395                 400

Pro Arg Ser Asp Pro Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln
                405                 410                 415

Pro Gln Ser Asp Pro Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln
            420                 425                 430

Pro Gln Ser Asp Ser Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln
        435                 440                 445

Thr Pro Ala Pro Ala Ala Ser Ser Val Pro Ser Pro Cys Asp Glu Ala
    450                 455                 460

Ser Pro Thr Pro Ser Ser His Pro Thr Pro Gly Ala Leu Glu Asp Pro
```

465 470 475 480

Ala Thr Pro Pro Ala Ser Glu Gly Glu Ser Pro Ser Ser Thr Pro Pro
485 490 495

Glu Ala Ala Pro Gly Ala Gly Pro Thr
500 505

<210> 93
<211> 1413
<212> DNA
<213> H.Sapiens

<400> 93
atggacacta ccatggaagc tgacctgggt gccactggcc acaggcccag cacagagctt 60

gatgatgagg actcctaccc ccaaggtggc tgggacacgg tcttcctggt ggccctgctg 120

ctccttgggc tgccagccaa tgggttgatg gcgtggctgg ccggctccca ggcccggcat 180

ggagctggca cgcgtctggc gctgctcctg ctcagcctgg ccctctctga cttcttgttc 240

ctggcagcag cggccttcca gatcctagag atccggcatg gggacactg gcggctgggg 300

acagctgcct gcccgttcta ctacttccta tggggcgtgt cctactcctc cggcctcttc 360

ctgctggccg ccctcagcct cgaccgctgc ctgctggcgc tgtgcccacc ctggtaccct 420

gggcaccgcc cagtccgcct gcccctctgg gtctgcgccg gtgtctgggt gctggccaca 480

ctcttcagcg tgccctggct ggtcttcccc gaggctgccg tctggtggta cgacctggtc 540

atctgcctgg acttctggga cagcgaggag ctgtcgctga ggatgctgga ggtcctgggg 600

ggcttcctgc ctttcctcct gctgctcgtc tgccacgtgc tcacccaggc cacagcctgt 660

cgcacctgcc accgccaaca gcagcccgca gccgccgggg gcttcgcccg tgtgcccagg 720

accattctgt cagcctatgt ggtcctgagg ctgccctacc agctggccca gctgctctac 780

ctggccttcc tgtgggacgt ctactctggc tacctgctct gggaggccct ggtctactcc 840

gactacctga tcctactcaa cagctgcctc agccccttcc tctgcctcat ggccagtgcc 900

gacctccgga ccctgctgcg ctccgtgctc tcgtccttcg cggcagctct ctgcgaggag 960

cggccgggca gcttcacgcc cactgagcca cagacccagc tagattctga gggtccaact 1020

ctgccagagc cgatggcaga ggcccagtca cagatggatc ctgtggccca gcctcaggtg 1080

aaccccacac tccagccacg atcggatccc acagctcagc cacagctgaa ccctacggcc 1140

cagccacagt cggatcccac agcccagcca cagctgaacc tcatggccca gccacagtca 1200

gactctgtgg cccagccaca ggcagacact aacgtccaga cccctgcacc tgctgccagt 1260

tctgtgccca gtccctgtga tgaagcttcc ccaaccccat cctcgcatcc taccccaggg 1320

gcccttgagg acccagccac acctcctgcc tctgaaggag aaagccccag cagcaccccg 1380

tcagaggcgg cccggggcgc aggccccacg tga 1413

<210> 94
<211> 419
<212> PRT
<213> H.Sapiens

<400> 94

Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro
1 5 10 15

Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp
20 25 30

Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro Ala Asn Gly
35 40 45

Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr
50 55 60

Arg Leu Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe
65 70 75 80

Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His
85 90 95

Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly
100 105 110

Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp
115 120 125

Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro
130 135 140

Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr
145 150 155 160

Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp
165 170 175

Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser
180 185 190

Leu Arg Met Leu Glu Val Leu Gly Gly Phe Leu Pro Phe Leu Leu Leu
195 200 205

Leu Val Cys His Val Leu Thr Gln Ala Thr Ala Cys Arg Thr Cys His
210 215 220

Arg Gln Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg
225 230 235 240

Thr Ile Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr Gln Leu Ala
245 250 255

Gln Leu Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu
260 265 270

Leu Trp Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser
275 280 285

Cys Leu Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr
290 295 300

Leu Leu Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Glu Glu
305 310 315 320

Arg Pro Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser
325 330 335

Glu Gly Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met
340 345 350

Asp Pro Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln Pro Arg Ser
355 360 365

Asp Pro Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser
370 375 380

Asp Pro Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser
385 390 395 400

Asp Ser Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala
405 410 415

Pro Ala Ala

<210> 95
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 95
ttcaagctt atggaatcat ctttctcatt tggagtgatc cttgctgtc 49

<210> 96
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 96
ttcactcgag ttagccatca aactctgagc tggagatagt gacgatgtg 49

<210> 97
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 97
gctcaaccca ctcatctatg cc 22

<210> 98
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 98
aaacttctct gcccttaccg tc 22

<210> 99
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 99
aaagcagcac cccgaatacc 20

<210> 100
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 100
catgatcaac ctgaacgtca c 21

<210> 101

<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 101
ttcaaagctt atggagtcgg ggctgctg 28

<210> 102
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 102
ttcactcgag tcagtctgca gccggttctg 30

<210> 103
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 103
gaatcctggc cgctatctgt gcactctacg 30

<210> 104
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 104
cctagagtgc acagatagcg gccaggatgc 30

<210> 105
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 105
aaccccatca tctacacgc 19

<210> 106
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 106
tgcctgtgga gccgctgg 18

<210> 107
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 107
gcataagctt ccatgtacaa cgggtcgtgc tgc 33

<210> 108
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 108
gcattctaga tcagtgccac tcaacaatgt ggg 33

<210> 109
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature

<223> Novel Sequence

<400> 109
gaagcccagc actgtttacc 20

<210> 110
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 110
tgacatacct gtccgcagcc 20

<210> 111
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 111
gatcaagctt atgacaggtg acttcccaag tatgc 35

<210> 112
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 112
gatcctcgag gctaacggca caaaacacaa ttcc 34

<210> 113
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 113
cagccccaaac atcccaagtc 19

<210> 114
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 114
[illegible]actta atccagcctc 19

<210> 115
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 115
gatcgaattc gcaggagcaa tgaaaatcag gaac 34

<210> 116
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 116
gatcgaattc ttatatatgt tcagaaaaca aattcatgg 39

<210> 117
<211> 20
<212> DNA
<213> Artificial Sequence

<400> 117
acagccccaa agccaaacac 20

<210> 118
<211> 22
<212> DNA
<213> Artificial Sequence

<400> 118
ccgcaggagc aatgaaaatc ag 22

<210> 119
<211> 19
<212> DNA
<213> Artificial Sequence

<400> 119
ctgaaagttg tcgctgacc 19

<210> 120
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 120
cgattatcca cactttgacc c 21

<210> 121
<211> 25
<212> DNA
<213> Artificial Sequence

<400> 121
gcataccatg aatgagccac tagac 25

<210> 122
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 122
ccatctcgag tcaagggttg tttgagtaac 30

<210> 123
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 123
ctgtctctct gtcctcttcc 20

<210> 124
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 124
gcaccgatct tcattgaatt tc 22

<210> 125
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 125
gcttcaaaca acttcatccc cc 22

<210> 126
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 126
acacacagca tagtagcg 18

<210> 127
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 127

cagagcttga tgatgaggac 20

<210> 128
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 128
cccataggaa gtagtagaag 20

<210> 129
<211> 9
<212> PRT
<213> Synthetic substrate peptide

<220>
<221> misc_feature
<223> Novel Sequence

<400> 129

Ala Pro Arg Thr Pro Gly Gly Arg Arg
1   5

<210> 130
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 130
gcgtaatacg actcactata gggagaccgc gtgtctgcta gactctattt cc 52

<210> 131
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 131
tgccacactg atgcaactcc 20

<210> 132
<211> 48
<212> DNA
<213> Artificial Sequence

<400> 132
gcgtaatacg actcactata gggagacctg ccacactgat gcaactcc 48

<210> 133
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 133
gcgtgtctgc tagactctat ttcc 24

<210> 134
<211> 50
<212> DNA
<213> Artificial Sequence

<400> 134
gcgtaatacg actcactata gggagaccgc acgcccctgt ttactatccc 50

<210> 135
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 135
gcaccaaaca caattccata agcc 24

<210> 136
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 136
gcgtaatacg actcactata gggagaccgc acaaacacca attccataag cc 52

<210> 137
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 137
gctacgcccc tctttactat ccc 23

<210> 138
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 138
gcgtaatacg actcactata gggagacctt atgagcagca attcatccc 49

<210> 139
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 139
cacacccacc aagaaatcag 20

<210> 140
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 140
gcgtaatacg actcactata gggagaccca cacccaccaa gaaatcag 48

<210> 141

<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 141
ttatgagcag caattcatcc c 21

<210> 142
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 142
gcgtaatacg actcactata gggagacccg attatccaca ctttgaccc 49

<210> 143
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 143
ctgaaagttg tcgctgacc 19

<210> 144
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 144
gcgtaatacg actcactata gggagaccct gctgaaagtt gtcgctgacc 50

<210> 145
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 145
cgattatcca cactttgacc c 21

<210> 146
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 146
gcgtaatacg actcactata gggagaccct gtaaaattca cacaagcacc 50

<210> 147
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 147
agaagacaga gcaacctcc 19

<210> 148
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 148
ggcgtaatac gactcactat agggagacca gaagacagag caacctcc 48

<210> 149
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature

<223> Novel Sequence

<400> 149
ctgtaaaatt cacacaagca cc 22

<210> 150
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 150
gcatggatcc tctttgctgt atttcaccct c 31

<210> 151
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 151
gcatgaattc acaatgccag tgataaggaa g 31

<210> 152
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 152
gatcaagctt ggaatgatgc cctttttgcca c 31

<210> 153
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 153
gatcctcgag catcattcaa agtaggtgg 29

<210> 154
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 154
gatcctcgag ctatgaactc aattccaaaa ataatttaca cc 42

<210> 155
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 155
gctacttgaa ctctacattt aatccaatgg tttatgcatt tttctatcc 49

<210> 156
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 156
ggatagaaaa atgcataaac cattggatta aatgtagagt tcaagtagc 49

<210> 157
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 157
gatcgaattc atggacacta ccatggaagc tgacc 35

<210> 158
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 158
gatcctcgag tcacgtgggg cctgcgcccg g 31

<210> 159
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 159
gcgtaatacg actcactata gggagaccgc gtgtctgcta gactctattt cc 52

<210> 160
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 160
tgccacactg atgcaactcc 20

<210> 161
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 161
gcgtaatacg actcactata gggagacctg ccacactgat gcaactcc 48

<210> 162
<211> 24

<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 162
gcgtgtctgc tagactctat ttcc 24

<210> 163
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 163
gcgtaatacg actcactata gggagacgcg acgccactct ttactatccc 50

<210> 164
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 164
gcacaaaaca caattccata agcc 24

<210> 165
<211> 52
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 165
gcgtaatacg actcactata gggagacgcg acaaaacaca attccataag cc 52

<210> 166
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 166
gctacgccac tctttactat ccc 23

<210> 167
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 167
gcgtaatacg actcactata gggagacctt atgggcagca attcatccc 49

<210> 168
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 168
cacacccacc aagaaatcag 20

<210> 169
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 169
gcgtaatacg actcactata gggagaccca cacccaccaa gaaatcag 48

<210> 170
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 170
ttatgagcag caattcatcc c 21

<210> 171
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 171
gcgtaatacg actcactata gggagacccg attatccacc ctttgaccc 49

<210> 172
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 172
ctgaaagttg tcgctgacc 19

<210> 173
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 173
gcgtaatacg actcactata gggagaccct gctgaaagtt gtcgctgacc 50

<210> 174
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 174

cgattatcca cactttgacc c 21

<210> 175
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 175
gcgtaatacg actcactata gggagaccct gtaaaattca cacaagcacc 50

<210> 176
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 176
agaagacaga gcaacctcc 19

<210> 177
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 177
gcgtaatacg actcactata gggagaccag aagacagagc aacctcc 47

<210> 178
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 178
ctgtaaaatt cacacaagca cc 22

<210> 179
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 179
gcatggatcc tctttgctgt atttcaccct c 31

<210> 180
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 180
gcatgaattc acaatgccag tgataaggaa g 31

<210> 181
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 181
acagccccaa agccaaacac 20

<210> 182
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 182
cccaagcagc aatgaaaatc ag 22

<210> 183
<211> 20
<212> DNA

<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 183
ctgtctctct gtcctcttcc 20

<210> 184
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 184
gcaccgatct tcattgactt tc 22

<210> 185
<211> 1188
<212> DNA
<213> H.Sapiens

<400> 185
aggctcgcgc ccgaagcaga gccatgagaa ccccagggtg cctggcgagc cgctagcgcc 60

atgggcccgg gcaaggccct gctggcccgt ctcctggtga tggtactggc cgtggcgctg 120

ctatccaacg cactggtgct gctttgttgc gcctaccgcg ctcagctccg cactcgagcc 180

tcagggctcc tcctggtgaa tctgtctctg ggccaactgc tgctgccgcc gctggacatc 240

cccttcacgc tgctcggtgt gatgcgcggg cggacaccgt cggcgccccg cgcatgccaa 300

gtcattggct tcctggacac cttcctggcg tccaacgcgg cgctgagcgt ggcggcgctg 360

agcgcagacc agtggctggc agtgggcttc ccactgcgct acgccggacg cctgcgaccg 420

cgctatgccg gcctgctgct gggctgtgcc tggggacagt cgctggcctt ctcaggcgct 480

gcacttggct gctcgtggct tggctacagc agcgccttcg cgtcctgttc gctgcgcctg 540

ccgcccgagc ctgagcgtcc gcgcttcgca gccttcaccg ccacgctcca tgccgtgggc 600

ttcgtgctgc cgctggcggt gctctgcctc acctcgctcc aggtgcaccg ggtggcacgc 660

agacactgcc agcgcatgga caccgtcacc atcaaggcgc tcgcgctgct cgccgacctg 720

caccccagtg tgcggcagcg ctgcctcatc cagcagaagc ggcgccgcca ccgcgccacc 780

aggaagattg gcattgctat tgcgaccttc ctcatctgct ttgccccgta tgtcatgacc 840

aggctggcgg agctcgtgcc cttcgtcacc gtgaacgccc agtggggcat cctcagcaag 900

tgcctgacct acagcaaggc ggtggccgac ccgttcacgt actctctgct ccgccggccg 960

ttccgccaag tcctggccgg cctggtgcac cggctgctga agagaacccc gcgcccagca 1020

tccacccatg acagctctct ggatctggcc ggcatggtgc accagctgct gaagagaacc 1080

ccgcgcccag cgtccacccа caacggctct gtggacacag agaatgattc ctgcctgcag 1140

cagacacact gagggcctgg cagggctcat cgccccсаcc ttctaaga 1188

<210> 186
<211> 363
<212> PRT
<213> H.Sapiens

<400> 186

Met Gly Pro Gly Glu Ala Leu Leu Ala Gly Leu Leu Val Met Val Leu
1 5 10 15

Ala Val Ala Leu Leu Ser Asn Ala Leu Val Leu Leu Cys Cys Ala Tyr
20 25 30

Ser Ala Glu Leu Arg Thr Arg Ala Ser Gly Val Leu Leu Val Asn Leu
35 40 45

Ser Leu Gly His Leu Leu Leu Ala Ala Leu Asp Met Pro Phe Thr Leu
50 55 60

Leu Gly Val Met Arg Gly Arg Thr Pro Ser Ala Pro Gly Ala Cys Gln
65 70 75 80

Val Ile Gly Phe Leu Asp Thr Phe Leu Ala Ser Asn Ala Ala Leu Ser
85 90 95

Val Ala Ala Leu Ser Ala Asp Gln Trp Leu Ala Val Gly Phe Pro Leu
100 105 110

Arg Tyr Ala Gly Arg Leu Arg Pro Arg Tyr Ala Gly Leu Leu Leu Gly
115 120 125

Cys Ala Trp Gly Gln Ser Leu Ala Phe Ser Gly Ala Ala Leu Gly Cys
130 135 140

Ser Trp Leu Gly Tyr Ser Ser Ala Phe Ala Ser Cys Ser Leu Arg Leu
145 150 155 160

Pro Pro Glu Pro Glu Arg Pro Arg Phe Ala Ala Phe Thr Ala Thr Leu
165 170 175

His Ala Val Gly Phe Val Leu Pro Leu Ala Val Leu Cys Leu Thr Ser
180 185 190

Leu Gln Val His Arg Val Ala Arg Arg His Cys Gln Arg Met Asp Thr
195 200 205

```
Val Thr Met Lys Ala Leu Ala Leu Leu Ala Asp Leu His Pro Ser Val
    210                 215                 220

Arg Gln Arg Cys Leu Ile Gln Gln Lys Arg Arg Arg His Arg Ala Thr
225                 230                 235                     240

Arg Lys Ile Gly Ile Ala Ile Ala Thr Phe Leu Ile Cys Phe Ala Pro
                245                 250                 255

Tyr Val Met Thr Arg Leu Ala Glu Leu Val Pro Phe Val Thr Val Asn
            260                 265                 270

Ala Gln Trp Gly Ile Leu Ser Lys Cys Leu Thr Tyr Ser Lys Ala Val
        275                 280                 285

Ala Asp Pro Phe Thr Tyr Ser Leu Leu Arg Arg Pro Phe Arg Gln Val
    290                 295                 300

Leu Ala Gly Met Val His Arg Leu Leu Lys Arg Thr Pro Arg Pro Ala
305                 310                 315                     320

Ser Thr His Asp Ser Ser Leu Asp Val Ala Gly Met Val His Gln Leu
                325                 330                 335

Leu Lys Arg Thr Pro Arg Pro Ala Ser Thr His Asn Gly Ser Val Asp
            340                 345                 350

Thr Glu Asn Asp Ser Cys Leu Gln Gln Thr His
        355                 360

<210>  187
<211>  29
<212>  DNA
<213>  Artificial Sequence

<220>
<221>  misc_feature
<223>  Novel Sequence


<400>  187
gcataagctt gccatgggcc ccggcgagg                                       29


<210>  188
<211>  28
<212>  DNA
<213>  Artificial Sequence

<220>
<221>  misc_feature
<223>  Novel Sequence


<400>  188
gcattctaga cctcagtgtg tctgctgc                                        28


<210>  189
```

<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 189
tgctgctttg ttgcgcctac 20

<210> 190
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> Novel Sequence

<400> 190
ttggacgcca ggaaggtg 18